Utilization of Computers for the Management and Development of Transportation Systems:

Papers and Proceedings of a U.S.-U.S.S.R. Seminar, Moscow 1975

Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9

Utilization of Computers for the Management and Development of Transportation Systems:

Papers and Proceedings of a U.S.-U.S.S.R. Seminar, Moscow

1975

Sydney Shulman, Administrative Editor

Translation and publication of this work was supported by grant No. GJ-41942 DCR 74-13801A02 from the National Science Foundation to the National Bureau of Economic Research, Inc. However, any opinions, findings, conclusions or recommendations expressed herein are those of the authors and do not necessarily reflect the views of the National Science Foundation or the National Bureau. Since the present volume is a record of conference proceedings, it has been exempted from the rules governing submission of manuscripts to, and critical review by, the Board of Directors of the National Bureau.

CONTENTS

Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9

INTRODUCTION

The U.S.-U.S.S.R. Conference on Utilization of Computers for the Management and Development of Transportation Systems took place in the Soviet Union on June 28-July 10, 1975. Fifteen papers were presented at the conference and are included in these proceedings. The purpose of the conference was to share concepts and techniques being used in the Soviet Union and the United States in the planning, design, implementation and use of computer models, simulations and programs for managing and operating transportation systems. There were open and frank discussions in regard to all the material presented. Both sides agreed that the next activity would be the development of mutually beneficial, longer term joint research.

Harvey McMains

AN OVERVIEW OF U.S. TRANSPORTATION: SIMILARITIES AND CONTRASTS WITH THE USSR

Holland Hunter *

This brief essay offers a primarily statistical sketch of the United States transportation system as it has evolved in the last quarter century or so. It is intended to set the stage for the research papers that follow, papers describing ways in which systems analysis aided by computers is being used in United States transportation. The sketch highlights some features of freight and passenger transportation that distinguish the United States situation from that of other countries, in particular that of the USSR. Both the U.S. and the USSR are large countries with large economies and large transportation systems. The contrasts, however, are as interesting as the similarities, and in our view much can be learned by examining both to see what they imply for the improved performance of each system.

Intercity Freight and Passenger Transportation

Over the last half century, United States intercity transportation has been marked primarily by a decline in the relative role of the railroads, matched by the steady rise of other carriers. For freight transportation, trends over the last thirty-five years are shown in table 1. Railroads accounted for two-thirds of the intercity freight traffic during the 1939-1948 decade; in recent years this share has fallen below 40%. Pipelines and intercity trucks have rapidly and steadily increased their share, each now accounting for over 22% of total intercity freight traffic. Freighters and barges on the Great Lakes, the Mississippi-Ohio-Missouri River system, and the coastal waterways of the Atlantic, Gulf, and Pacific Coasts have steadily raised the volume of their traffic and

* Professor, Haverford College

Table 1

Domestic Intercity Freight Traffic, U.S., by Carrier, Five-Year Totals, 1939-43 through 1969-73, in Billions of Metric Ton-Kilometers

	Railroads	Motor Vehicles	Inland Waterways	Oil Pipelines	Airways	Total
1939-43	4070	457	912	520	-	5989
1944-48	5202	621	1032	848	1	7704
1949-53	4631	1295	1177	1057	3	8163
1954-58	4513	1745	1507	1527	4	9296
1959-63	4438	2197	1583	1723	7	9948
1964-68	5354	2746	1993	2424	16	12524
1969-73	5742	3264	2386	3319	26	14737
			Percent Shares			
1939-43	68.0	7.6	15.7	8.7	-	100
1944-48	67.5	8.1	13.4	11.0	-	100
1949-53	56.7	15.9	14.4	13.0	-	100
1954-58	48.6	18.8	16.2	16.4	-	100
1959-63	44.6	22.1	15.9	17.3	.1	100
1964-68	42.7	21.9	15.9	19.4	.1	100
1969-73	39.0	22.1	16.2	22.5	.2	100

Source: Compiled from annual data in ton-miles (1 metric ton - km = 1.46 short ton-mile) in U.S. Bureau of the Census, Historical Stat. of the U.S. (1960), Series Q1-11 (adjusted for coverage changes); Statistical Abstract of the U.S: 1974, p. 547; and U.S. Interstate Commerce Commission, 88th Annual Report, p. 120.

maintained their share of about one-sixth of the total.

United States intercity passenger travel is heavily dominated by the passenger automobile, as shown in table 2. Since 1950, privately owned automobiles have accounted for at least 86% of all intercity passenger travel; their share in 1960 reached 90% but since then has fallen back slightly. The balance is handled by airways, buses, railroads, and internal waterways, with the air share steadily rising to over 10% and the share of buses and railroads shrinking. The absolute volume of bus travel has been fairly steady, but rail passenger travel has fallen by two-thirds over the last quarter century. Today, air transport is clearly the dominant form of public intercity passenger transport, accounting for well over 70% of intercity passenger travel not done by automobile.

The revenues earned by carriers for all their domestic freight and passenger service are displayed in table 3. Trends reflect the physical changes already noted but with significant distinctions. The railroad share of revenue, for example, is less than their share of freight, since railroads carry a good deal of low-value freight. The truck share of revenues, conversely, is higher than their share of freight, and, recently, intercity trucking has taken in more than 42% of all transportation revenues. The airline share of total passenger and freight revenues has grown rapidly from 4% to 20%. Pipeline revenues, though they have tripled over the last quarter century, are still only 3% of all transportation revenue since costs and charges for this form of movement are extremely low. Low per-unit costs also characterize internal waterway freight traffic, where the revenue share has been falling though the volume of traffic continues to grow and the freight share is stable.

Table 2

Domestic Intercity Passenger Taffic, U.S., by Carrier, Selected Years, 1950-1973, in billions of Passenger-kilometers

	Private Automobiles	Airways	Buses	Railroads	Inland Waterways	Total
1950	704.9	16.1	41.8	51.5	1.9	816.2
1955	1025.1	37.0	40.2	46.7	2.7	1151.7
1960	1136.2	54.7	30.6	35.4	4.3	1261.2
1965	1316.4	93.3	38.6	29.0	5.0	1482.3
1970	1651.2	191.5	40.2	17.7	6.4	1907.0
1973						

Percent Shares

	Private Automobiles	Airways	Buses	Railroads	Inland Waterways	Total
1950	86.4	2.0	5.1	6.3	0.2	100
1955	89.0	3.2	3.5	4.1	0.2	100
1960	90.1	4.3	2.4	2.8	0.4	100
1965	88.8	6.3	2.6	2.0	0.3	100
1970	86.6	10.0	2.1	0.9	0.4	100
1973						

Source: Passenger-mile data from U.S. Dept. of Commerce, Bureau of the Census, Statistical Abstract of the U.S., 1975, p. 562.

Table 3

Revenues From Domestic Freight and Passenger Traffic, U.S., by Carrier, Selected Years, 1950-1973, in Millions of Dollars

	Railroads	Trucks	Airlines	Pipelines	Buses	Waterlines	Total
1950	10,147	3,737	558	442	539	330	15,753
1955	10,831	5,535	1215	678	552	452	19,263
1960	10,203	7,214	2129	770	667	427	21,410
1965	11,054	10,068	3609	904	885	426	26,946
1970	12,824	14,585	7131	1188	1062	502	37,292
1973	15,864	20,800	9605	1446	1135	615	49,465
				Percent Shares			
1950	64.4	23.7	3.6	2.8	3.4	2.1	100
1955	56.2	28.7	6.3	3.5	2.9	2.4	100
1960	47.7	33.7	9.9	3.6	3.1	2.0	100
1965	41.0	37.4	13.4	3.3	3.3	1.6	100
1970	34.4	39.1	19.1	3.2	2.9	1.3	100
1973	32.1	42.1	19.4	2.9	2.3	1.2	100

Source: Statistical Abstract of the U.S., 1975, p. 560.

Table 4 presents summary data for 1973, showing not only the ton-kilometers of freight handled by each of the four major carriers but also their tons originated and the resulting quotient showing each carrier's average length of haul. Intercity freight typically moves for long distances in the United States, reflecting the continental dimensions of the economy. Railroads and pipelines have an average length of haul above the overall national average, while waterways and highways show a somewhat lower figure. These crude averages, both in aggregate and for the individual modes, summarize an underlying distribution including a great deal of short-haul movement offset by significant amounts of extremely long-haul traffic. Strictly local movement by trucks in urban areas is specifically excluded from these statistics.

The massive expansion of automotive passenger and freight transportation is dramatically symbolized by the data in table 5 showing the route length of railroads and highways in the United States over the period 1890-1973. Initially, the total length of the railroad network was greater than the total length of surfaced highways, but by 1920 the highway network exceeded the railroad system in length. After 1930, U.S. railroads began a gradual abandonment of little-used roadway; the 1973 length of the railroad system is below that of 1910. In sharp contrast, however, the length of surfaced highways almost quadrupled from 1930 to 1973. It will be seen that the total length of all highways outside towns and cities has grown only modestly; the basic change has involved upgrading through providing existing roads with all-weather surfaces. These summary figures also understate the qualitative improvement that has come with modern highway construction, especially with the 69,000 kilometers of the Interstate Highway System. Intercity passenger

Table 4

Freight Shipments and Average Length of Haul, U.S., 1973, by Carrier, in Kilometers and millions of Metric Tons

	Ton-kilometers (billions)	Tons Originated (millions)	Average Haul (kilometers)
Railroads	1,253	1,466	855
Highways	737	1,833	402 (a)
Waterways	523	902	580
Pipelines	740	1,091	678
Four-carrier Total	3,253	5,292	615

(a) Class I and II common and contract carriers, 1971. See ICC, Trans. Stat. of the U.S., part 7, release 2, pp. 48, 54, 164.

Source: U.S. Dept. of Commerce, Bureau of the Census, Stat. Abstract of the U.S., 1975, pp. 580-81, 586, 596.

Table 5

Length of Railroads and Highways, U.S., Selected Years, 1890-1973, in Thousands of Kilometers

		Highways (outside towns and cities)		
	Railroads	Surfaced	Unsurfaced	Total
1890	251	NA	NA	NA
1900	311	248 (a)	NA	NA
1910	388	328	NA	NA
1920	418	594	4113	4707 (b)
1930	418	1117	3725	4842
1940	396	2156	2956	4812
1950	381	2702	2110	4812
1960	370	3484	1531	5015
1970	354	3880	1220	5100
1973	348	3920	1191	5111

(a) 1904
(b) 1921

Source: U.S. Dept. of Commerce, Bureau of the Census, Historical Stat. of the U.S. to 1957, pp. 429, 458; Stat. Abstract of the U.S., 1975, pp. 564, 581.

and freight movement by highway has been substantially enhanced by this national grid of limited-access, divided-lane superhighways. Of course, at the same time, these improvements in the highway system have been occurring, corresponding technological improvements have occurred in railroading (e.g., centralized traffic control and improvements in the geometry and weight of trackage), so that even though railroad system kilometers have contracted, the total capability of the railroad system has remained at least constant and probably enlarged.

The usage of roads and highways outside cities in the United States is divided roughly three to one between passenger automobiles and trucks, with buses accounting for a negligible share of total vehicle kilometers. The figures in table 6 show that, measured in billions of vehicle-kilometers per year, the volume of use by passenger cars in 1973 was 3.5 times as large as in 1940, bus use had doubled, and truck use was five times as great. This growth in usage went well beyond the expansion in nonurban highway mileage, so that intercity traffic density per route kilometer more than doubled. As a crude overall average, in 1940 one could see thirteen vehicles per hour pass one point on the network every hour every day through the year. By 1973 this figure had risen to twenty-eight. The truck share of highway usage was gradually expanding at the expense of buses and passenger cars and, though the growth of bus usage appeared to be leveling off, growth in truck and passenger car use of the highways was vigorous through 1973. Higher fuel prices may check this growth somewhat, but the forces pressing for further growth seem very strong.

TABLE 6

INTERCITY MOTOR-VEHICLE MOVEMENT AND DENSITY,
U.S., SELECTED YEARS, 1940-1973

(in billions of vehicle-kilometers and hourly vehicle-kilometers per kilometer of surfaced highway)

	Passenger Cars	Buses	Trucks	Total	Hourly Number of Vehicles
1940	193.9	2.3	48.6	244.8	13
1950	291.4	3.4	91.4	386.2	16
1960	488.1	3.7	131.5	623.3	20
1970	654.0	4.5	215.8	874.3	26
1973	715.0	4.7	247.8	967.5	28
		Percent Shares			
1940	79.2	0.9	19.9	100	
1950	75.4	0.9	23.7	100	
1960	78.2	0.6	21.2	100	
1970	74.8	0.5	24.7	100	
1973	73.9	0.5	25.6	100	

Source: Derived from U.S. Department of Commerce, Bureau of the Census, Stat. Abstract '75, pp. 571 and 564

Urban Transportation

In the last forty years urban passenger transportation in the United States has shifted markedly away from public carriers toward the private automobile. Table 7 shows the number of passengers carried by the major means of mass transit at various times between 1930 and 1973. On subways and elevated electric trains, the number dropped from 13 billion annually to less than 2 billion. Railroad commutation passengers dropped from 435 to 183 million per year. Passengers using buses increased markedly from 1930 to 1950, but their number has since been cut in half. The trolley bus similarly showed a rapid expansion up to 1950 and an equally sharp decline since then. Overall, the annual number of passenger trips made on these four major means of mass transit has dropped from 17.5 billion in 1950 to less than 7 billion in 1973. Much public policy has been directed toward promoting the use of mass transit, but the United States urban population has stubbornly continued to prefer the private passenger automobile.

The census of 1970 asked a 15% sample of urban and rural households about how working members of each household got to work. Their answers are summarized in table 8. Two-thirds of the labor force rode to work in automobiles that they drove, and more than 10% rode as passengers. Those who walked to work made up 7.4%, and 9% used bus, street car, subway, elevated train, railroad, or taxicab. A few used other means, and about 4% of the urban labor force (over 12% of the rural labor force) worked at home. This pattern reflects in part the way that residences and work places have been dispersed in postwar U.S. metropolitan regions; it is increasingly inconvenient or even impossible to get from home to work and back by means other than the passenger automobile.

TABLE 7

URBAN PASSENGER TRAFFIC, U.S., BY PUBLIC CARRIER, SELECTED YEARS, 1930-1973

(in millions of passengers carried)

	Subway and Elevated Electric RR	Railroad Commutation	Motor Bus	Trolley Bus	Total
1930	13,072	435	2,479	16	16,002
1940	8,325	229	4,239	534	13,327
1950	6,168	277	9,420	1,658	17,523
1960	2,313	203	6,425	657	9,598
1970	2,116	206	5,034	182	7,538
1973	1,921	183	4,642	97	6,843

Sources: U.S. Dept. of Commerce, Bureau of the Census, Hist. Stat. of the U.S., p. 464; Stat. Abstract '75, p. 579; and (for railroads)

Assoc. of Am Railroads, Railroad Trans., A Statistical Record 1921-63 (Wash., 1965), p. 24; Stat. of Railroads of Class I (Aug. 1974), p. 7.

TABLE 8

MEANS OF TRANSPORTATION TO WORK, U.S., 1970
URBAN AND RURAL WORKERS

(in millions of workers 16 years old or over)

	Urban	Rural	Total
Passenger automobile			
As driver	38,133	12,565	50,698
As passenger	6,702	2,322	9,024
On foot	4,506	1,183	5,689
Bus or streetcar	4,116	129	4,245
Subway or elevated train	1,762	6	1,768
Railroad	473	29	502
Taxicab	278	18	296
Other, or worked at home	2,258	2,371	4,629
Total	58,228	18,623	76,851

Percent Shares

	Urban	Rural	Total
Passenger automobile			
As driver	65.5	67.5	66.0
As passenger	11.5	12.5	11.7
On foot	7.7	6.3	7.4
Bus or streetcar	7.1	0.7	5.5
Subway or elevated train	3.0	--	2.3
Railroad	0.8	0.2	0.7
Taxicab	0.5	0.1	0.4
Other, or worked at home	3.9	12.7	6.0
Total	100.0	100.0	100.0

Source: U. S. Department of Commerce, Bureau of the Census, Stat. Abstract '75 p. 578

The streets and highways of U.S. cities and towns are dominated by passenger automobiles, as shown in the data of table 9. Automatic counting devices and inspectors at selected locations regularly measure the number of vehicles of each type using the roads. Passenger cars accounted for 86% of all vehicle-kilometers in 1940, and their 84% share in 1973 was only slightly lower. The truck share rose from 13% to 16%, while the share of buses dropped from slightly under 1% to only 0.3%. The absolute volume of bus movement has stopped growing. Though the total length of urban streets and highways has expanded, traffic density (measured as the hourly number of vehicles passing a given point) has roughly doubled since 1940.

Some Basic Dimensions of Soviet Transportation

The statistics above were presented in metric tons and kilometers to facilitate comparisons with analogous Soviet measures of transportation. Here a few Soviet statistics are presented to illustrate some basic contrasts between the transportation systems of the two economies. Table 10, similar in form to table 4, shows for 1973 the level of activity of the four major freight carriers. It is immediately apparent that railroads dominate Soviet freight transportation to a greater extent than has been true in the United States for many decades. Soviet railroads carry more than twice as much freight as U.S. railroads for approximately the same average distance. The Soviet figure for truck traffic relates mainly to local rather than intercity movement (the average haul in the USSR is 16 kilometers compared to 402), so for comparability most of this Soviet truck traffic should be excluded from a comparison of intercity traffic. Conversely, a portion of Soviet

TABLE 9

URBAN MOTOR-VEHICLE MOVEMENT AND DENSITY,
U.S., SELECTED YEARS, 1940-1973

(in billions of vehicle-kilometers and hourly vehicle-kilometers per kilometer of surfaced highway)

	Passenger Cars	Buses	Trucks	Total	Hourly Number of Vehicles
1940	207.8	1.9	31.7	241.4	N.A.
1950	293.7	3.2	54.4	351.3	77
1960	458.3	3.4	71.9	533.6	88
1970	795.8	3.5	129.7	929.0	117
1973	953.0	3.4	182.0	1,138.4	128
			Percent Shares		
1940	86.1	0.8	13.1	100	
1950	83.6	0.9	15.5	100	
1960	85.9	0.6	13.5	100	
1970	85.7	0.4	13.9	100	
1973	83.7	0.3	16.0	100	

Source: Derived from U. S. Dept. of Commerce, Bureau of the Census, Stat. Abstract '75, pp. 571 and 564.

Table 10

Freight Traffic, Shipments, and Average Length of Haul, USSR, 1973, by Carrier, in Billions of Metric Ton-Kilometers, Millions of Metric Tons, and Kilometers

	Ton-Kilometers (Billions)	Tons Originated (Millions)	Average Haul (Kilometers)
Railroads	2,958	3,346	884
Highways	284	18,244	16
Waterways	190	419	453
Pipelines	439	421	1,043
Four-Carrier Total	3,871	22,430	173

Source: TsSU, Narkhoz SSSR 1973, pp. 503, 504, 510, 516-17

maritime freight traffic is domestic, moving between Soviet ports, and should be included, though it is not covered in table 10.

Soviet internal waterways are not as favorably located (e.g., the Volga River system does not directly connect Moscow with the eastern Ukraine or the Urals, nor is there a good water connection between the Urals and western Siberia) as those in the United States, and the volume of Soviet waterways freight traffic is thus only about a third of the U.S. level. Soviet pipelines, though less well developed than U.S. pipelines, are already moving almost 40% as much as U.S. pipelines over longer average distances. The aggregate volume of Soviet freight traffic, using comparable coverage, would probably be some 10%-15% greater than U.S. intercity freight traffic. This aggregate measure is hard to interpret, since it covers differing commodity structures in the two countries' freight traffic, reflecting in turn the differences between the two economies in their output composition.

Soviet intercity passenger traffic is only about one-third as great as passenger movement in the United States. The difference lies mainly, of course, in the absence of appreciable intercity travel by passenger automobile (it is not yet even estimated in Soviet transportation statistics). In marked contrast to U.S. railroads, Soviet railroads are still carrying an increasing volume of intercity passenger traffic; in 1973 they handled almost six times as many passenger-kilometers as U.S. railroads handled in 1950. Soviet bus traffic, similarly, in 1973 was six times the 1950 U.S. level. Soviet air traffic has been growing very rapidly and, in 1973, reached about 40% of the 1973 U.S. level. Modest movement on inland waterways is about the same in both economies. While this is not the place to examine them in detail,

TABLE 11

DOMESTIC INTERCITY PASSENGER TRAFFIC, USSR,
BY CARRIER, SELECTED YEARS, 1950-1973

(in billions of passenger-kilometers)

	Railroads	Buses	Airways	Inland Waterways	Sea	Total
1950	88.0	5.2	1.2	2.7	1.2	98.3
1955	141.4	20.9	2.8	3.6	1.5	170.2
1960	170.8	61.0	12.1	4.3	1.3	249.5
1965	201.6	120.5	38.1	4.9	1.5	366.6
1970	265.4	202.5	78.2	5.4	1.6	553.1
1973	296.6	253.9	98.8	5.9	1.9	657.1
			Percent Shares			
1950	89.5	5.3	1.2	2.8	1.2	100
1955	83.1	12.3	1.6	2.1	0.9	100
1960	68.5	24.5	4.8	1.7	0.5	100
1965	55.0	32.9	10.4	1.3	0.4	100
1970	48.0	36.6	14.1	1.0	0.3	100
1973	45.2	38.6	15.0	0.9	0.3	100

Source: TsSU, Narkhoz '73, p. 500

major issues of national policy surround the question of whether the relative absence of intercity passenger automobile movement in the USSR can, or should, continue in the future.

Table 12 shows the length of railroads and highways in the USSR in a format similar to that of table 5 above. One notes that, though the Soviet railroad network is very large, it has only about one-third the length of the U.S. system before the latter began contracting. Most observers would say that the U.S. system was overbuilt, but perhaps this crude comparison suggests that additional railroad mileage will be useful for the Soviet economy. As to highways, the contrast is most striking: in 1973, the U.S. had 3.6 times as long a road system as the USSR, and for surfaced highways the ratio was 6.6 to 1. Russian and Soviet writers have been commenting on the problem of "roadlessness" in Russia for at least a century and, clearly, much remains to be done. Foreign visitors riding along the skeletal existing paved highway system are unlikely to be aware of the lack of paved roads throughout most of the country.

The data on Soviet urban passenger traffic in table 13 show how rapidly it has increased during the postwar period. The total number of passengers carried per year has risen from 9 to 53 billion, almost all moved by traditional carriers. The growth constrasts sharply with the U.S. decline shown in table 7 (the latter, of course, confined to the minority who do not travel by passenger automobile). In the USSR since 1950, the number of passengers carried by autobus has risen very rapidly, as has the much smaller number who use taxis. In 1950, streetcars dominated the scene, but their volume of traffic has been declining since 1965 and now accounts for only 15% of the total. Trolley bus, subway, and railway commutation traffic has steadily increased though not to the

TABLE 12

LENGTH OF RAILROADS AND HIGHWAYS, USSR,
SELECTED YEARS, 1940-1973

(in thousands of kilometers)

	Railroads	Highways Surfaced	Highways Unsurfaced	Total
1940	106	143	1,388	1,531
1950	117	177	1,373	1,550
1960	126	271	1,095	1,366
1970	135	512	852	1,364
1973	137	598	800	1,398

Sources: TsSU, Transport i Sviaz SSSR (1972). pp.89, 262

TABLE 13

URBAN PASSENGER TRAFFIC, USSR, BY CARRIER
SELECTED YEARS, 1950-1973

(in millions of passengers carried)

	Auto Bus	Street car	Trolley Bus	Subway	Railway	Taxi	Total
1950	1,001	5,157	945	629	955	43	8,730
1955	4,294	6,367	1,858	937	1,392	45	14,893
1960	10,797	7,842	3,055	1,148	1,713	389	24,944
1965	17,771	8,242	4,298	1,652	2,049	717	34,729
1970	25,901	7,962	6,122	2,294	2,616	1,144	46,039
1973	30,458	7,998	7,298	2,727	2,970	1,461	52,912
				Percent Shares			
1950	11.5	59.1	10.8	7.2	10.9	0.5	100
1955	28.8	42.8	12.5	6.3	9.3	0.3	100
1960	43.3	31.4	12.2	4.6	6.9	1.6	100
1965	51.2	23.7	12.4	4.7	5.9	2.1	100
1970	56.2	17.3	13.3	5.0	5.7	2.5	100
1973	57.6	15.1	13.8	5.1	5.6	2.8	100

Sources: TsSu, Transport i Sviaz, (1972) pp. 99-100, 244, 246-47, 256-57;
Narkhoz '73, pp. 504, 522-23, 525,
Narkhoz '58, p. 588 (for autobus 1950 and 1955).

same extent as bus traffic. These summary data are not conclusive, but they suggest that urban public transit carriers in the USSR have a record of healthy growth.

Analytic Contrasts Between United States and Soviet Transportation.

Behind these statistics for the two continental economies lie several fundamental contrasts in the conditions facing transportation agencies and the users of their services. It is worth noting them briefly for the light they shed on the problems that systems analysis can deal with. The most important concerns relations between supply and demand.

In the United States, the demand for transportation confronts an ample supply of transportation capacity. Almost everywhere, in fact, increments of transportation demand could be quickly accommodated by at least one carrier, and rival transportation agencies are eager to handle more traffic. In the USSR, however, transport demand presses hard on transport supply, both freight and passenger, and added demand is not easily accommodated. The Soviet government did not inherit an over-built railroad system, nor has it put major resources into a well-developed system of local, district, and interregional paved highways. Soviet transport capacity has barely kept up with burgeoning demand.

In the United States, the typical producing enterprise can choose among multiple sources of supply for its inputs, multiple routes linking the supplier with the enterprise, and multiple carriers to provide the freight service. The availability of alternatives gives the enterprise considerable leverage in obtaining good service. Prompt and flexible adjustments are facilitated since bottlenecks at particular points can

be bypassed. Speed and reliability of service are encouraged. There may be costs imposed on the overall economy by excess capacity in the transportation sector, not continuously and fully utilized, but these costs may be more than offset by benefits to the overall economy in reduced inventories, faster production, improved product quality, and speedier adaptation to customer needs.

Several of the following research papers focus on factors that enter into the economic decisions made by shippers under these conditions. United States transportation agencies need accurate information on what their customers want, since without this information the carrier may not obtain and ratain the shipper's orders. The factors that enter into travelers' choices among modes of travel and specific passenger carriers are the object of similar analysis. Railroads, airlines, bus lines, and trucking firms in the United States are not free to set the conditions under which they offer service, solely on the basis of their own convenience and internal cost-minimizing considerations. The choices made by shippers and travelers have a decisive influence on the growth and welfare of the transport sector.

Other papers concentrate on ways that systems analysis is used by United States transportation agencies to minimize their operating costs and provide efficient service. Mr. Johnson's paper shows how railroads use computers to keep track of freight cars and how trucking firms match trucks with shipment flows. Mr. Ehrlich's paper shows how Eastern Airlines uses an elegant systems analysis to find a cost-minimizing solution to the problem of assigning aircraft and flight crews to a pattern of flights offered the public, a pattern which has already been carefully tailored to fit the evidence on what the public desires.

The delegation of United States transportation specialists was impressed in July 1975 by what they heard and saw concerning the ways that Soviet transportation agencies are using systems analysis and computers. Clearly, simlar problems lead to similar answers. Soviet analysts are searching for cost-minimizing solutions in the internal operations of the carrier. Railroads plan their heavy freight traffic flows, and the Moscow subway carefully monitors its heavy passenger traffic. Aeroflot uses its ticketing and reservation service to minimize empty-seat-kilometers, and the maritime fleet keeps track of its farflung merchant marine. Intracity Soviet truck traffic is centrally organized in interesting ways that go beyond the decentralized operations of United States trucking firms, suggesting that Soviet methods of trucking might be a fruitful topic for joint analysis.

In spite of the differences between the USSR and the United States in balances of supply and demand for freight and passenger transportation, it seems clear that a great many problems of efficient operation for each means of transportation can usefully be studied by very similar methods of computer-aided systems analysis. Here an exchange of operating experiences and a comparison of analytic methods can be illuminating for experts on both sides. In addition, it seems to us that further study of the factors influencing the demand for transportation, both in the minds of shippers and in the minds of travelers, can serve the interests of the Soviet economy even though intercarrier competition is absent. Long-run growth and efficiency in any economy require that economic activity should respond correctly to the objective needs of consumers and producers, and for the transportation sector this means that demand-sensitive analysis is necessary, no matter who own the means of production.

FREIGHT DEMAND MODELLING: A POLICY SENSITIVE APPROACH

Mark N. Terziev, Moshe Ben-Akiva, and Paul O. Roberts *

1. INTRODUCTION

Determining the volume of cargo that will flow in a given freight market is the starting point for any quantitative analysis of freight policy. This is true whether the issues being addressed are those of a carrier who would like to consider changes in a particular pricing policy or a government attempting to justify the major capital expenditure for a new facility. In spite of what appears to be an obvious need for analysis tools, very little work has been done to provide such a capability in the freight area.

One may account for this situation in several ways. First, there has not been data available to use in model development. This is not a condition that should be allowed to continue to exist since the data could be obtained once the needs are known precisely. Second, there has not been a clearly articulated statement of purpose for such a model. This is understandable in the light of rising government interest in the problems of the freight sector. Finally, the lack of an adequate theoretical framework has hindered both the collection of data and the statement of purpose for the appropriate analytical tools. This lack is one which we intend to address directly in this writing.

Demand Supply Equilibrium in Freight Markets

First, it is appropriate to attempt to clear up the second deficiency, the lack of a clearly stated purpose. This can best be done by putting the demand forecasting element into perspective relative to the demand-supply equilibrium of freight markets. The shipper of freight makes a number of decisions which influence the size and nature of the

* Director, Center for Transportation Studies and Professor, M.I.T.

market between any two points. It is useful to take the viewpoint of the shipper in constructing the demand element of an equilibrium model of the system. The attributes of the commodity being shipped and the characteristics of the shipper and the markets in which he is engaged will be important, as will information concerning the transportation system available to the shipper in making his decisions.

These characteristics of the transportation system are influenced directly by the decisions made by the carriers competing in a given freight market. The size and type of equipment, its scheduling, reliability, etc., all determine directly, along with the tariff, the level of service which the shipper will receive and ultimately influence his demand for service. The supply element of the system treats the decisions made by the carrier in his offerings.

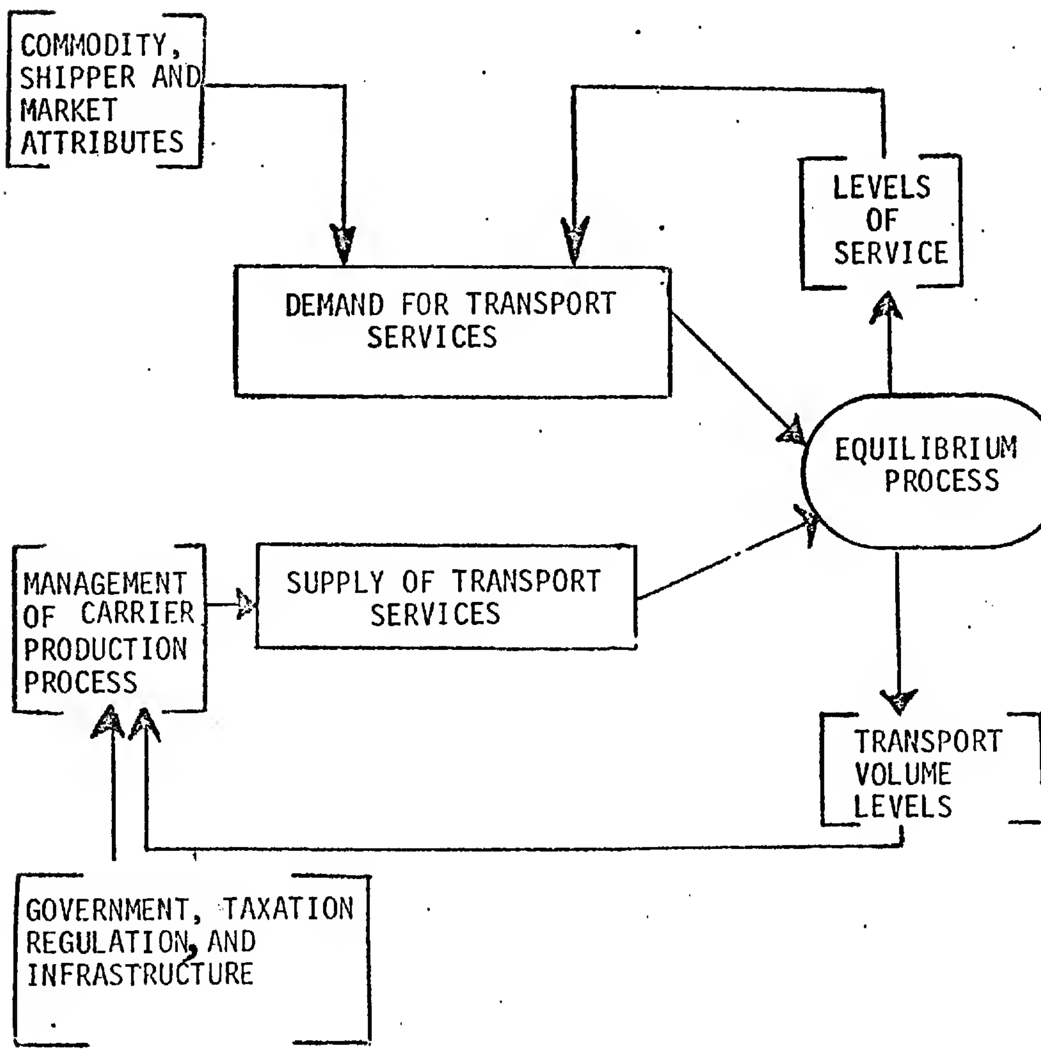
Equilibrium occurs in the system because of the range of choice that exists and the feedback nature of the information between supply and demand. Thus, as the tariff for transport by one mode increases relative to alternative modes there is diversion to the lower cost modes by those shippers most sensitive to shipment costs. The output of the equilibrium process includes the volumes of shipments by the different modes of freight transport and the levels of service experienced by the shippers at these shipment levels. (See figure 1.1.)

The Role of Government in the Process

The process is influenced by government in a variety of ways, all on the supply side. Government typically controls entry, mergers, ownership, finances, rates, and routes by means of the regulatory agencies. Through the executive and administrative agencies taxes, tax credits,

FIGURE 1.1

NATURE OF THE EQUILIBRIUM PROCESS



tolls, subsidies, environmental and safety standards are set and transport infrastructure is provided. In those efforts government is attempting to act in a manner that is equitable and in the best interests of its constituency, within a rather broad interpretation of the meaning of this phrase.

Making the Models Policy Sensitive

For any set of models to be policy sensitive they must incorporate the variables reflected in the policies of interest. Thus, if the policy involves the impact of a change in equipment on the competitive advantage of a particular mode, the supply models must be able to represent these changes in the level of service as they will eventually be perceived by the shipper. If the policy involves a change in pricing, the level of service vector must reflect the new tariff as seen by the shipper. A variety of supply side models may be needed to explore the full range of policies of interest since the control variables employed directly by the government and the carrier are on the supply side.

On the demand side the different policies are represented by the list of characteristics making up the level of service vector. If a level of service vector can be specified which captures the principal determinants of travel demand on the part of the shipper, then a demand model employing this level of service vector can be used to handle a large range of policy alternatives. Each of the policy alternatives to be studied must be transformed by the supply models into the variables in this vector. As a consequence, one well-specified demand model can handle a variety of supply side policy investigations. Of course, specifying the variables to be incorporated in the level of service vector will be an important undertaking.

Role of Each Element in the Equilibrium Process

Each of the elements of the system has a different function. The demand model attempts to replicate the decision-making behavior of an individual shipper of a commodity faced with a number of alternatives. This behavior varies with the attributes of the product being shipped, the shipper, and the market. The supply models must transform the policy variables under study into a level of service vector for each alternative which might be chosen. Both are involved in the equilibrium process since the level of service inputs into the demand element may not elicit travel demand volumes which are compatible with the original level of service estimates produced by the supply element. The volume must be consistent with the level of service at equilibrium.

Our focus here is on the demand element since it is the least developed and the most crucial in some sense. That is, once a policy to be investigated has been transformed into the level of service attributes using even the roughest of supply models the volume levels can be obtained. Even if the transformation of the policy into level of service attributes is done on a completely intuitive basis, the model can be useful. Therefore, it is necessary to consider the possible range of policy questions which might be addressed.

Possible Users of Such a Model

A fundamental understanding of the way in which freight flows are determined would be of use to government planners, policy-makers and regulators, as well as carriers and shippers. An analysis of freight flows is required for the evaluation of government policy options in the following areas of interest:

1. Freight transportation facility planning in various corridors
2. Economic analysis of projects involving the modification or expansion of the freight transportation systems
3. Impact of changes in regulation which affect the price and level of service of freight carriage
4. Overall modal policy planning

Carriers would also benefit from a better understanding of how their supply policies interact with the shippers' demand for the transport of goods. The major areas of concern to the carrier include the following:

1. Analysis of potential markets
2. Impacts of changes in the level of service
3. Impacts of changes in pricing policy
4. Evaluation of changes in regulations
5. Determining the feasibility of new services

Shippers would like to coordinate their short-run policies with their long-run plans. However, the complex interaction of the shippers' behavior with the policies of the other parties makes the task of coordination very difficult. A better knowledge of this interaction process would help shippers deal with the following issues:

1. Evaluation of location options
2. Determination of capacity allocation and production plans

Possible Policy Issues to be Investigated with the Model

Perhaps the most important set of issues facing transportation policy-makers today are those dealing with the future of the railroads. How can they be made more competitive? Are there ways to rationalize the network with line abandonments, mergers, or government ownership of

the right-of-ways? Would new, advanced, multimodal services be successful? If so, how should the services be priced? How should tariffs be established? What impact would new services have on conventional service? For any of these to be handled in more than a superficial manner, a demand model of the type described here is necessary.

Another important set of policy questions is concerned with deregulation. The analysis of deregulation policies can be undertaken on a market-by-market basis with the techniques outlined here. To address the questions at a more macroscopic level it will be necessary to place the demand model into a more aggregate framework and consider the system-wide impacts. This model, however, represents the necessary first step in such a process.

Policies involving fuel conservation and air quality have recently received attention. The demand element is especially important in this area since the ability to predict the impact of various policies on the overall level of demand may be crucial. The extent of the impact on choice of mode or shipment size of various transportation control strategies may also be of concern.

Many other questions such as truck size-weight issues, expansion of the interstate system, rate absorption in ocean shipping, the imposition of tolls on inland waterways, the development of the domestic air freight system, etc., can be addressed. The key is the translation of the policy questions into level of service vectors for each of the alternatives under investigation, using either a formally developed set of supply models or good intuitive judgement as previously suggested. The demand model is therefore a crucial element in any quantitative analysis whether the policies to be investigated are those of a private sector entrepreneur,

a public sector decision maker, or a legislative body seeking the impact of proposed changes in legislation.

The Need for Data on Shipper Behavior

At the present time, forecasting freight movements is not widely done. A good bit of freight movement data exists, but it is hard to coordinate and use the data. Their volume alone makes it difficult to deal with. This would not be such a problem if the data did not lack several crucial elements needed for accurate forecasting. The data currently available are largely historical flows, by commodity, over given corridors and specific facilities. The response to government or carrier policy is by the shipper. Thus, the shipper's views are extremely important. To be useful, the data must capture the behavior of the shipper. In making his choice, the shipper views the alternatives that are available to him and makes his choice between them based on the conditions of transport and the relative costs. Data concerning the choices made by the shipper rarely exist.

It would be difficult to use the data on shipper behavior if it were available only in tabular form. It is necessary to capture the values of the shipper; for example, how he weighs transit time as opposed to reliability or cost. This can be most easily done by the use of a carefully structured econometric model which captures and makes available the "valued" choices of the shipper. This is not to say that these are not useful in their own right. Quite the contrary, there are a wide variety of uses for the data, including the calibration of other models for other uses. At the moment, however, data on the choice behavior of shippers is unavailable in a coordinated and useful form.

Why Policy-Sensitive Freight Demand Models Don't Currently Exist

Part of the reason the data does not exist and why models which incorporate it also do not exist is because the models of the past were not behavioral models. Instead they were aggregate models which captured correlations of aggregate quantities rather than individual shipper behavior. The development of disaggregate behavioral models using modern methods of econometrics is fairly recent. The development of data sets which can be used to calibrate or test them has not yet been accomplished.

New methods for predicting passenger travel demand have advanced very rapidly over the past few years. The profession has come a long way since the invention of the "gravity model." Disaggregate behavioral econometric models employing less data collection and more efficient data utilization have been developed (CRA, 1972; Ben Akiva, 1973).

The disaggregate approach appears to be equally attractive for freight demand models. In fact, preliminary investigations have shown that disaggreagte freight demand models are both workable and encouraging. The impact of this new modelling approach is likely to have considerable more import than the various aggregate econometric models developed during the past ten years. The pioneer efforts were illustrative, but not very practical for prediction purposes.

The Role of Data Collection

The design of such a model must proceed in conjunction with the design of a data scheme and methods for data collection. Disaggregate models require much less data for calibration than aggregate models and, therefore, it is now entirely practical to gather the data needed for a freight demand model.

Objective of Current Work

The overall objective of this research effort is to build upon the existing work that has been done in freight demand modeling by employing the recent developments in disaggregate passenger travel demand models. The first steps in this direction are to review existing sources of data (including disaggregate data sets if they are available) and to survey the literature on freight demand models with particular emphasis on previous research done on disaggregate models. The major elements of this research effort are the specification of a system of policy-sensitive disaggregate freight demand models based on a realistic theory of shipper's choice of origin, destination, frequency, shipment size, and mode of transport, and the specification of data requirements and their collection methods. Ultimately, this methodological phase should be followed by actual data collection, model estimation, sensitivity tests and validation, and, finally, the use of these models as policy analysis tools.

2. FRAMEWORK FOR FREIGHT DEMAND MODELLING

In this section the framework for freight demand modelling is described in terms of the specific shippers' choices and the variety of factors that determine the pattern of commodity flows.

Shipper Behavior

The pattern of freight flows results from the actions of numerous actors. As noted previously, the basic decision-making unit in determining the demand for freight transport is the individual shipper. Depending upon the ownership of the freight, the decision-maker is usually located at either the origin or the destination of a shipment. The behavior

that leads to a demand for freight transport can best be described from the point of view of this individual decision-maker.

A strategic decision (and the most long-range action of any shipper) is a choice of location. For intercity freight the location of production plants and warehouses substantially determines the requirements for, and the characteristics of, available freight transport services.[1]

For a given set of production and consumption points, the pattern of commodity flows is determined by the individual shipper's selection of markets and suppliers. The quantity of freight of commodity type k that the producer at point i is shipping to his warehouse at destination j, for example, is determined by the shipper's choice to market his product in a specific volume in the market area served by the warehouse at point j.

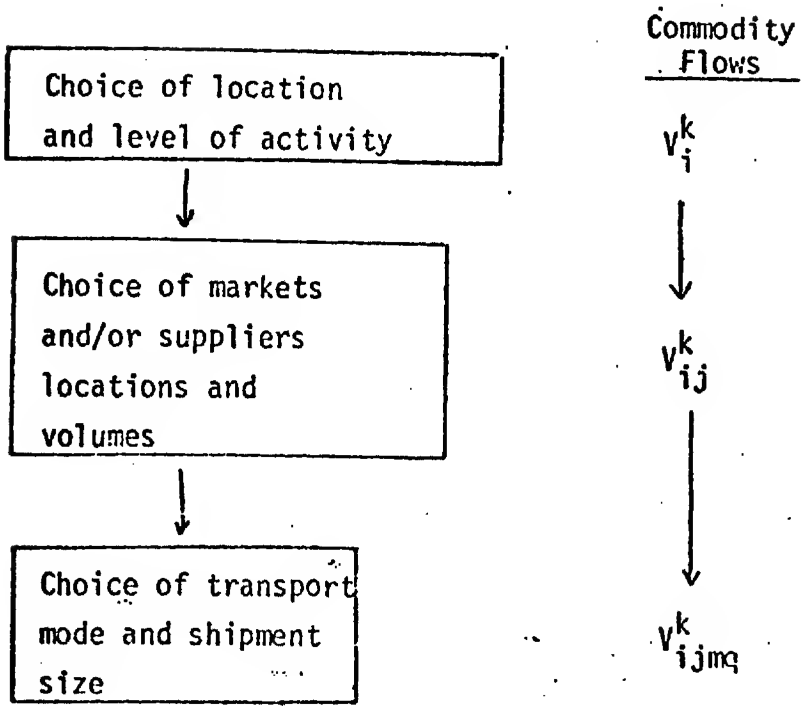
Given the quantity of commodity k shipped from origin i to destination j over a certain time period, an individual shipment (as represented by a waybill) is determined by the choices of mode m and shipment size q. There are a variety of feasible combinations of modes and shipment sizes that can be used to transport a given annual volume between two locations. Note that the annual volume and the shipment size also determine the frequency at which shipments are made.

The Choice Hierarchy

Thus, the basic unit of commodity flow--the individual shipment--is determined by a complex hierarchy of choices made by an individual shipper. This hierarchy is depicted in figure 2.1. The sequence that is assumed in this hierarchy represents the different time lags in a shipper's response to a change in transport policy. For some shippers

FIGURE 2.1

HIERARCHY OF SHIPPER CHOICES



and commodity types, the transport mode and shipment size may be altered on very short notice and at a very small cost. However, a location change almost always requires significant time and capital expenditures.

It is important to note that this choice hierarchy does not imply a long-run, sequential, causal relationship. In a long-range decision, a lower-level choice such as the choice of the transport mode, for example, influences the choice of location, and vice versa.

This means that a lower-level choice is determined in two ways: strategic (long run) and tactical (short run). A change in a higher-level choice is likely to result in a reconsideration of all lower-level choices. Thus, in a strategic decision all the choices are jointly determined, and in a tactical decision some high-level choices are fixed and only lower-level choices are adjusted.

Another important implication of this hierarchy is that at any given point in time we are likely to observe a shipper at a disequilibrium point. There is a threshold level below which level of service changes do not influence shippers' choice in the short run. For this reason, we can observe an inefficient pattern of commodity flows which are due to the cost involved in adjusting the choices high on the choice hierarchy.

Factors Affecting Shipper Behavior

The freight transportation situation at the level of an individual shipper can be described in terms of four groups of characteristics. First, there are shipper attributes which include those characteristics relating to the location of the plant, the location of supplies, and long-run operating policies. Second, there are market attributes which

are related to both the long-run choice of markets and the short-term equilibrium demand for the commodity being shipped. Third, there are commodity attributes which include the physical characteristics of the good as well as information concerning its handling and use. The fourth group consists of transport level of service attributes which depend not only on the mode, but also the commodity being shipped and the shipment size being considered.

In the long run, the choice of location, the level of activity, the shipment size, and the choice of mode are variable at the discretion of the shipper. The commodity being shipped is assumed fixed, and a unique commodity type is associated with each basic decision-making unit.

Commodity Attributes

The list of commodity attributes that influence the demand for freight transport might include the following:

1. Value per pound--basic commodity value at the origin
2. Shipping density--both of the basic commodity and of the commodity when packed for shipping
3. Shelf life--number of days to spoilage or obsolesence
4. Product use--manufacturing, processing, final consumption
5. Method of inventory control--single-item, multi-item
6. Stockout consequence--no cost, contribution loss, probability increase with days, shutdown

A shipper would also consider the temporary storage facilities associated with the commodity being shipped, although this is variable over the long run.

Shipper Attributes

Some shipper attributes which are fixed in the short run are variable in the long run and are determined by the strategic decisions at the top of the choice hierarchy. The choice of plant location and the choice of suppliers and markets fixes the distances over which each commodity must be transported. The choice of plant size (level of activity) brackets the range in the volume of each commodity which must be moved. These considerations are summarized by the following variables:

1. Annual production volume--amount produced
2. Number of establishments--number of production plants
3. Location of establishment--on rail siding, water access, etc.
4. Ownership--whether producer owns the freight
5. Price at origin--whether price is established at factory
6. Decision-maker--point where inventory decisions are made

A shipper would also consider the use of private carriage. This would effectively replace the short-run mode choice decision with a long-run investment decision.

Market Attributes

Market attributes reflect the market demand for the commodity being shipped. For the most part, the market equilibrium is variable over the short run but it can be assumed to be external to the freight demand decision. Therefore, market attributes influence the intermediate decisions relating to the level or volume of shipment. The market variables that influence shippers include:

1. Sales in consuming industry--amount consumed

2. Size of consuming industry--number of firms consuming
3. Characteristics of consuming population--socioeconomic characteristics
4. Size of consuming population--number
5. Location of establishment--on rail siding, etc.
6. Ownership--whether supplier owns the freight
7. Price at destination--whether price is established C.I.F. destination
8. Decision-maker--point where inventory decisions are made

Mode/Size Choices

In the short run the shipper is left with only the choices of mode and shipment size. The mode shipment size options considered by a shipper include all or some of the following:

1. Freight forwarder
2. Airfreight parcels
3. Air container
4. Private truck
5. Contract carrier
6. Common carrier truck LTL
7. Common carrier truck FTL
8. TOFC Plans I-V
9. Rail FCL
10. Rail unit train
11. Inland barge
12. Barge container
13. Sea container
14. Pipeline

In some cases, these tactical options will be precluded by higher order strategic decisions. For example, the choice of plant location, supplier location, and market location will determine whether rail, barge, air, and pipeline are available options in the mode choice decision.

Level of Service Attributes

Once the list of possible mode/shipment size alternatives has been edited down to the available short-run options, then the final decision will be based on the modal level of service attributes. For each available option this group of characteristics contains:

1. Waiting time
2. Travel time
3. Time reliability
4. Loss and damage
5. Transport tariff
6. Other costs
7. Minimum shipment size
8. Special services

The factors influencing shippers' behavior are summarized in figure 2.2. Here an overall view can be gained of the interrelationships between the variables which are important in the process. We now turn to an examination of the existing literature to determine the extent to which models employing these variables in the appropriate way already exist in useful form.

3. LITERATURE REVIEW

A number of studies of freight demand have appeared in the transportation and economics literature during the past ten years. The demand

FIGURE 2.2

THEORETICAL FRAMEWORK FOR A MODEL OF FREIGHT DEMAND

$$V^k_{ijmq} = f\ (T, C, M, S)$$

V = quantity of freight
k = commodity type
i = origin
j = destination
q = shipment size
m = mode

T = Transport Level of Service Attributes	C = Commodity Attributes	M = Market Attributes	S = Shipper Attributes
Waiting time	Value/lb.	Sales-consuming industries	Annual production volume
Travel time	Density	Size-consuming industries	Number of establishments
Time Reliability	Shelf life	Population characteristics	Location of establishment
Loss and damage	Ownership	Size-consuming population	Ownership
Transport tariff	Product use	Location of establishment	Price at origin
Other costs	Inventory system	Ownership	Decision-maker
Minimum shipment size	Pricing policy	Price at destination	
Special services		Decision-maker	

models developed in these studies generally fall into two categories and representative models from both groups are discussed in this section. The first group consists of very simple, macroscopic models which predict only the total freight flow between all origins and destinations. The data required to calibrate this type of model is relatively easy to find and therefore most of these models have been empirically estimated. The principal use of these models is for the prediction of the impact of system-wide changes in prices or the level of service.

The second group of models consists of microscopic demand models which address the problem of predicting the freight flow between origin-destination pairs. These models include variables relating to specific operations. Therefore, these models can be used to analyze policy operations dealing with operations on particular links in the freight transportation network. However, very few of the microscopic demand models have been thoroughly tested, although some preliminary empirical work has been done.

The analysis of different freight transportation policy issues requires models with different levels of detail. Therefore, a freight demand model may not necessarily be applicable to all policy evaluation tasks. Both micro and macro models are useful tools.

Sloss (1971)

Sloss has presented a model for predicting the total volume of intercity freight traffic carried by trucks. The dependent variable in his model is the annual tonnage of goods transported by for-hire trucks on intercity hauls (denoted V_t). The independent variables are: the average truck revenue per ton (C_t), the average rail revenue per ton (C_r), and

a variable representing the general level of economic activity (E). The economic activity variable was defined as the unweighted sum of farm cash income, the value of building permits issued and the value of shipments of manufactured goods. These variables were combined in a product form model which was calibrated using regression analysis. The general form of the model is:

$$V_t = a_0 \, C_t^{a_1} \, C_r^{a_2} \, E^{a_3}$$

The data for the model were obtained from annual reports published by the Dominion Bureau of Statistics in Ottawa, Canada. Sloss's empirical results imply that the volume of truck traffic is directly related to the price of rail service and inversely related to the price of truck service. Sloss also concluded that the quantity of truck traffic is positively correlated with the general level of economic activity. These results seem intuitively reasonable.

Sloss's model could be used for a macro-level evaluation of alternative pricing strategies. However, it could not be used to evaluate other system-wide options dealing with changes in operating speeds, improvements in reliability, etc. Nor could this model be used to predict the impact of a policy dealing with changes in the price of carriage of a particular commodity between a particular origin and destination.

A. D. Little (1974)

As part of an analysis of domestic waterborne shipping, A. D. Little, Inc. has developed a model which predicts the percentage of intercity freight traffic moved by water carriers. Unlike most other models which

have appeared in the literature, this model emphasizes the strategic (long-run) aspects of the mode choice decision. In this model, the level of service variables such as time and cost have been replaced by variables which indicate the compatibility between the shipper's operations and the operations of waterborne carriers. For example, the following variables have been used to characterize shippers in a BEA region:[2]

1. Total flow--used because ships and barges are best suited for carrying high volumes

2. Value per ton--used because water modes typically carry low-value goods

3. Bulk commodity indicator--used because ships and barges are well suited for the loading and unloading of bulk goods

4. Seasonality--used because the high fixed cost of dock facilities makes it desirable to maintain a high volume of usage throughout the year

It was found that these four variables are highly correlated with the fraction of plants located on bodies of water in BEA regions.

No direct measures of level of service were included in this model. However, two proxy variables were included to represent the differences between the barge and truck (and rail) levels of service as they might be perceived in the long run. These variables are the highway distance (approximately the same as the rail distance) and the ratio of the water route distance to the highway distance.

The variables described above were used in a nonlinear equation to predict the fraction of traffic moved by water carriers. The data used in this study were records of flows of fifteen commodity groups between BEA regions.

In terms of policy analysis, strategic models of this kind can be quite useful. First, they can help identify key variables to use as long-run policy instruments. Second, they point out how the inherent advantages of a mode interact with shipper attributes and market attributes in the long-run decision-making process. However, it is obvious that these models cannot be used for the evaluation of specific pricing and operating policies which affect long-run, as well as short-run, shipper behavior.

Perle (1964)

Perle conducted one of the first major studies of freight demand. Perle used product form models with prices as the only indication of the level of service on each of the models. He calibrated one model for the total volume of truck freight and a second for the total volume of rail freight. The independent variables in both models are the average rail rate (C_r) and the average truck rate (C_t). Perle used dummy variables to represent the influence of certain commodity groups and certain regional characteristics. Since he was using time series data, he also included a dummy variable to represent the year. But he did not use any market attributes or variables related to the size and state of the economy. The general forms of the models are:

$$V_r = a_r C_r^{b_{rr}} C_t^{b_{rt}} \left(\pi_i \, e_{ri}^{R_i} \; \pi_j \, d_{rj}^{T_j} \; \pi_k \, f_{rk}^{Z_k} \right)$$

$$V_t = a_t C_r^{b_{tr}} C_t^{b_{tt}} \left(\pi_i \, e_{ti}^{R_i} \; \pi_j \, d_{tj}^{T_j} \; \pi_k \, f_{tk}^{Z_k} \right)$$

where:

R_i = 1 if observation is in ith region

= 0 otherwise

T_j = 1 if observation is in jth year

= 0 otherwise

Z_k = 1 if observation is in kth commodity group

= 0 otherwise

Perle used ICC reports to find the total tons carried by each mode during several years. This data was then used to calibrate several variations of the basic model shown above.

On the whole, Perle's models did not fit well. Several of the estimated coefficients had the wrong sign. However, the dummy variables representing commodity specific effects and regional effects significantly improved the fit of the models. The dummy variable representing the year was less useful. These results indicate that further effort should be made to explicitly include commodity and market attributes in freight demand models.

In terms of policy analysis, the uses and shortcomings of Perle's models are similar to those of the Sloss model. His models can be used for the evaluation of pricing options which would affect an entire system. But Perle's models cannot relate price changes on particular hauls to impacts on traffic in particular market segments. Moreover, they neglect many other policy variables such as travel time and reliability. Nevertheless, Perle's work is an important contribution to the field of freight demand analysis. He used a careful, methodological approach and his results demonstrate the importance of commodity and market attributes.

Mathematica (1967)

Mathematica did some theoretical work in freight demand analysis as part of the Northeast Corridor Transportation Project. One of the models which they proposed is an adaptation of the Quandt/Baumol abstract mode model used in passenger demand modeling.[3] The general form of the model is:

$$V_{ijm} = a_0\, P_i^{a_1}\, P_j^{a_2}\, Y_i^{a_3}\, Y_j^{a_4}\, M_i^{a_5}\, M_j^{a_6}\, N_{ij}^{a_7}\, (T_{ij}^b)^{b_0} (T_{ij}^r)^{b_1} (C_{ij}^b)^{d_0} (C_{ijm}^r)^{d_1}$$

where:

V_{ijm} = volume of freight flow from i to j by mode m

P_i, P_j = population of the origin and destination

Y_i, Y_j = gross regional product of the origin and destination

M_i, M_j = industrial character indices such as the percent of the labor force employed in mining and manufacturing

T_{ij}^b = least shipping time from i to j

T_{ijm}^r = travel time by mode m divided by the least travel time from i to j

C_{ij}^b = least cost of shipping from i to j

C_{ijm}^b = cost of mode m divided by the least cost from i to j

N_{ij} = number of modes serving i and j

Unlike the other models discussed above, this model deals with the flow of goods between particular origin-destination pairs. Furthermore, this model contains a greater range of market attributes and level of service attributes, although it does neglect a significant number of policy variables, such as time reliability.

A second flaw in the Quandt/Baumol model involves intermodal effects. In this model, changes in the cost and time on an inferior mode will have

no effect on the volume carried by the "best" model. This is certainly counter-intuitive. It is not readily apparent how the model could be modified to avoid this problem.

Mathematica (1967)

Mathematica has also proposed a freight demand model completely unlike those discussed above. This model was developed through the application of classical microeconomics and inventory theory. Without going through the derivation of the model, the general line of development is:

1. Express total annual variable transportation costs of the industry in city j which is receiving commodity k from city i as the sum of the direct shipping costs, total in-transit carrying costs, and safety inventory costs
2. Differentiate total cost to get marginal cost and apply the optimality rule that marginal cost equals marginal revenue

Under some assumptions about the cost functions and some approximations to make the algebra work, the following demand function was derived:

$$V^k_{ijm} = a_0 + a_1 (\nabla P^k_{ij}) + a_2 C^k_{ijm} + a_3 (S^k_j T^k_{ijm}) + a_4 (S^k_j + T^k_{ijm})^{1/2}$$

where:

V^k_{ijm} = volume of commodity k moving from i to j by mode m

∇P^k_{ij} = the difference in value of commodity k between city i and city j

S^k_j = time between shipments of commodity k to j

T^k_{ijm} = average travel time of commodity k from i to j on mode m

This kind of microeconomic, inventory-based model might be a good policy evaluation tool for a shipper doing an analysis of his location options and general operating policy. It may not be obvious from the general form of the model, but the following important factors are implicitly represented by the variables and parameters:

1. interest rate on capital
2. value of the commodity
3. price elasticity of demand for the commodity
4. warehousing costs
5. rate of sales of the commodity
6. cost of stocking-out of the commodity
7. ordering costs

The derivation of the model from a microeconomic point of view allows the inclusion of these important factors. However, the transition from a model of an individual firm to a model of the aggregate behavior of all firms in the origin city was not made explicit in the Mathematica report. It is not entirely clear how the aggregation issue should be handled for disaggregate models of this type.

<u>Antle and Haynes (1971)</u>

Antle and Haynes have suggested still another approach to freight demand analysis. They used a linear discriminant function to predict the choice of mode, given some basic information about the shipment. The general form of their model is:

$$\hat{Z} = a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + a_5X_5 + a_6X_6 + a_7X_7$$

where:

X_1 = annual tonnage shipped

X_2 = distance by principal mode

X_3 = average travel time by principal mode

X_4 = average shipment size

X_5 = rate on principal mode

X_6 = rate on competing mode

X_7 = handling cost on principal mode

These seven independent variables are used to characterize a single shipping decision. To use the model, the values for a principal mode and a competing mode are substituted into the model and a value of Z is calculated. This value is compared with a test value Z to predict whether the principal or the competing mode will be chosen. If the distribution of each of the independent variables is known for a given commodity, for a given origin-destination city pair, then the volume V^k_{ijm} can be estimated. So, even though this model is based on discriminant theory, it can be used to produce the same kind of demand information as the other models discussed in this section.

The Antle and Haynes model was calibrated on a small amount of data that was collected on the shipment of coal, coke, petroleum, and chemicals by barge and by rail in the Upper Ohio River Valley. Using just the section of the data on coal shipments, they estimated a model with only four significant independent variables: travel time and rate on the principal mode, shipment size, and handling cost. Although this model fitted the data reasonably well, it should be regarded with some suspicion because it lacks so many important variables. The failure of the rate on the competing mode to appear in the calibrated model was

particularly disappointing.

This model could be used to study the volume of coal carried by each mode on a particular link in the freight transportation network. However, the model contains no commodity attributes, which means that a separate model would have to be calibrated for each commodity being studied if the model is to be used to evaluate policies related to particular goods. Despite its shortcomings, the Antle and Haynes study is important because it helped pave the way for the application of stochastic, disaggregate qualitative choice models to freight demand.

Kullman (1973)

Kullman used a binary logit model with aggregate data to predict the truck-rail modal split as a function of level of service attributes, commodity attributes, and market attributes. The general form of his model is:

$$\left(\frac{\text{Rail Share}}{\text{Truck Share}}\right)^{\text{Commodity } k}_{i \text{ to } j} = f\text{ (distance, annual volume, value per ton, difference in rail and truck freight rates, difference in rail and truck transit time, difference in rail and truck transit time reliability)}$$

As a policy tool, Kullman's model has several desirable characteristics. It includes a number of level of service variables, including reliability. It also includes annual volume consumed, which is a market attribute. Furthermore, Kullman's model is to some extent commodity abstract because it includes commodity value. (See p. 38, last paragraph.) Although this model does neglect some important factors, it is a relatively sophisticated example of a microlevel model.

As a tool for microlevel analysis Kullman's model has three shortcomings. In the first place, it should include a wider range of commodity and market attributes as well as some shipper attributes. Secondly, Kullman's empirical work has shown that several important levels of service variables (such as travel time reliability) cannot be represented by aggregate data. Aggregate reliability data contains very little variation among observations and, therefore, the effect of reliability on mode choice cannot be determined. Third, since it is only a modal split model, in order to predict V^k_{ijm}, an external estimate of the total volume of commodity k moving from i to j must be provided. The major contribution of Kullman's model is the inclusion of a wider range of variables than most other models. It also demonstrates the feasibility of using the logit model in freight demand analysis.

Hartwig and Linton (1974)

The work of Hartwig and Linton is the most recent effort in the area of disaggregate freight demand modeling. They used the logit, probit, and discriminant functions to model the individual shipper's mode choice between full-load truck and full-load rail.

The data base used in this study consisted of 1213 freight waybills for full-load truck and rail shipments of consumer durables. From the bills they determined:

1. The origin and destination of the shipment from which they determined mileage
2. The date shipped and the date received from which they determined transit time
3. The freight rate and total charge
4. Shipment weight

5. The type of commodity being shipped from which they found the value per unit of the shipment

In their models, Hartwig and Linton used the difference in truck and rail transit times, the difference in truck and rail freight charges, the difference in truck and rail reliability (as measured by the difference in the standard deviation of the transit time distributions) and the value of the commodity as the independent variables. The dependent variable as in the Antel and Haynes and Kullman models is the mode choice probability. Using these variables, they experimented with logit, probit, and discriminant models. In general, they found that both logit and probit performed well, while the discriminant approach yielded somewhat poorer results.

The Hartwig and Linton study indicates that the disaggregate logit model can be used very successfully in freight demand modelling. However, their model needs to be expanded if it is to be used as a microlevel policy tool. Market attributes and shipper attributes should be added along with some additional level of service variables. Furthermore, the addition of commodity attributes would make the model useful to analyze a wider range of commodity types and differential process policies.

In summary, a number of models have been developed and tested which can be used in macrolevel policy analysis. To date, these models have not been used extensively. More detailed models capable of predicting flows of particular commodities between particular cities have developed more slowly. The principal reason for the slow progress in this area is the scarcity of data at the required level of detail. Nevertheless, this second type of model is a potentially powerful method for performing a detailed analysis of intercity freight operations. It is important to

note that, given a microlevel demand model, an appropriate model for macrolevel analysis can in principle be developed by aggregating the micro model over the distribution of its independent variables. This aggregation of microlevel demand models is discussed in the following section.

4. DISAGGREGATE MODELLING OF SHIPPER BEHAVIOR

This section presents a brief description of the methodology suggested for modelling shipper behavior. The approach outlined here avoids many of the problems presented in the work which has been done to date as described in the previous section. The recommended approach involves the use of a disaggregate behavioral model.

Why Disaggregate Models?

A freight demand model is most likely to be used in a study of freight facilities and policies or for a market analysis of freight movements of a certain type. The forecasts required are not for the individual shipper's freight demand but rather for some aggregate patterns of freight movements. Since the basic decision-making unit for freight demand is an individual shipper, it follows that the aggregate demand is simply a sum of individual shipper's demands.

The starting point for a freight demand model formulation is a theory of the behavior of an individual shipper. However, for the aggregate forecasting of freight demand, the model of individual shipper's behavior must be aggregated. The aggregation is based on groupings of shippers by their geographical locations, industry type, etc. If data is available at the level of the individual shipper, then this aggregation can be performed either before or after model estimation.

An aggregate model is estimated, using the means or totals of the variables included in the model for an aggregate group of shippers. Sometimes parameters of the distributions of the independent variables within the aggregate groups are also included. A disaggregate model is estimated, using the values of the variables for individual shippers. The aggregation for forecasting is then performed by integrating the disaggregate demand model over the distribution of the independent variables' values for the aggregate group of shippers.

The advantages of estimating a disaggregate model are numerous. One of the most important points is their efficient use of data. Since the data is not aggregated prior to model estimation, less original data is required to obtain reliable estimates of the model's coefficients. Aggregate data is characterized by a significant loss in the variability of important variables. Therefore, estimation of an aggregate model often fails to produce reliable coefficients for variables such as reliability, waiting time, etc., which do not vary substantially between large geographic zones but have considerable variations within these zones.

Because a disaggregate model is not estimated using averages for specific aggregate groupings, it is not tied to a specific area or a specific aggregation scheme. Once estimated, a disaggregate model can be used for a wide range of applications at different areas and different levels of detail without a need for model reestimation.

Disaggregate freight mode choice models have already been estimated on a trial basis by Hartwig and Linton (1974) and by Antle and Haynes (1971). These studies have proven that a disaggregate modelling approach is workable and attractive for freight demand modelling.

Modelling Qualitative Choice[4]

At the level of the individual shipper (or the single shipment), the decisions with respect to transport mode and shipment size, for example, are characterized as a selection from a set of qualitative alternatives. In qualitative choice, the alternatives can be indexed but cannot be rank ordered. Locational, modal, and shipment-size choices are characterized as qualitative choice problems.

Consider an individual shipper making a selection from a set of alternatives A_t where i and j denote alternatives in the set. For example, the set of alternatives for a model of mode choice and shipment size, includes all the feasible combinations of available modes and shipment sizes. For a model that includes a destination choice as well, the set of alternatives includes all combinations of modes, sizes, and destinations. Each alternative in the choice set is associated with some measure of attractiveness (cost) which the shipper is attempting to maximize (minimize). Given that one and only one alternative is selected from the choice set then alternative i will be selected if and only if:

$$U_{it} \geq U_{jt}, \forall j \in A_t$$

where U_{it} denotes the attractiveness of alternative i to shipper t. In a mode choice problem, for example, this implies that mode i will be selected if and only if its attractiveness is greater than or equal to the attractiveness of any other mode that is available.

The attractiveness of alternative i is a function of the attributes of the alternative (e.g., transit time, rate, loss and damage, consumption rate at destination, etc.). Two types of attributes directly

influence short-run freight shipment decisions: transport modal attributes (or level of service attributes) and market attributes. However, the same transport alternative will be evaluated differently by different shippers of different commodities. Thus, the attractiveness of a given freight transport alternative is expressed as a function of shipper and commodity attributes as well as the attributes of the alternatives. Given data on the observed behavior of shippers and the values of the important attributes, the attractiveness functions must be estimated if we are to be able to forecast freight demand under varying conditions.

Due to measurement errors, unobservable information, and other deficiencies in the available data, we are, in general, unable to estimate these attractiveness functions with certainty. Therefore, the deterministic choice criterion will not always correctly predict the actual choices. We can express the attractiveness of an alternative as consisting of two parts, observable and random, as follows:

$$U_{it} = u_{it} + \varepsilon_{it} \qquad (1)$$

where u_{it} is the observable part of the attractiveness function and ε_{it} is an observable random element. If the attractiveness measures of the alternatives include random components, then only the probability of choosing each option can be predicted, rather than a deterministic choice.

The probability that alternative i will be selected by shipper t equals the probability that the attractiveness (cost) of alternative i is greater or smaller than, or equal to, the attractiveness (cost) of all other alternatives that are available. Formally, this statement can be expressed as follows:

$$P(i:A_t) = \text{Prob } [U_{it} \geq U_{jt}, \quad j \in A_t] \tag{2}$$

where $P(i:A_t)$ is the probability of shipper t selecting alternative i from his choice set A_t. Substituting equation (1) in (2) we get:

$$P(i:A_t) = \text{Prob } [\varepsilon_{jt} - \varepsilon_{it} \leq u_{it} - u_{jt}, \quad j \in A_t] \tag{3}$$

This expression implies that the joint probability distribution of the random components determines the form of the model that relates the systematic attractiveness functions to the choice probabilities.

One specific assumption about the random elements joint distribution leads to the multinomial logit model which is the only probabilistic choice model that has been extensively applied to multiple choice problems. The random elements are assumed to be independently and identically distributed as follows:

$$P(\varepsilon \leq w) = e^{-ne^{-w}}$$

where ne is anh positive constant substituting this distribution assumption in equation (2) and integrating results in the logit model which is expressed as follows:

$$P(i:A_t) = \frac{e^{u_{it}}}{\sum_{j \varepsilon A_t} e^{u_{jt}}}$$

The systematic attractiveness functions are usually restricted to be linear with the parameters:

$$\begin{aligned} u_{it} &= X_{it}\ \Theta \\ &= \sum_{k=1}^{K} X_{itk} \cdot \Theta_k \end{aligned}$$

where X_{it} is a K*1 vector of independent variables describing alternative i and shipper t and Θ is a K*1 vector of coefficients that must be estimated.

The independent variables can be formulated in this model as alternative specific or as generic. In the first case, a variable such as transit time, for example, assumed different coefficients for different alternatives, e.g., alternative modes. For a generic variable formulation the differences among alternatives are only due to differences in the variable values. The variable of transit time, for example, has the same coefficients for all modes and only the value of the variable itself differs. A formulation in which all the independent variables are generic is also called an "alternative abstract" model and can be used to predict the demand for "new" alternatives, e.g., new mode.

The data required to estimate the model are: the attributes of the alternatives (i.e., transport and market attributes), the shipper and commodity attributes, and the actual choice made. Note that the data is required for the attributes of the chosen as well as the nonchosen alternatives. The observed dependent variable takes a value of one if the alternative was chosen and zero otherwise. However, the forecasts produced by the model are choice probabilities for the alternatives. The forecasted probabilities satisfy the following conditions:

$$< P(i:A_t) < 1, \quad i \in A_t$$

$$\text{and} \sum_{i \in A_t} P(i:A_t) = 1$$

The estimation technique which is often used is the maximum likelihood method. There are no limitations on the number of variables or on

the number of alternatives. Furthermore, the number of alternatives need not be identical for all observations.

Abstract Commodity Representation

One issue of significant importance which remains to be addressed concerns the number of models to be developed. For example, if one model must be formulated, calibrated, and tested for each commodity, then the effort is doomed to failure at the outset. There are so many possible commodities and subcommodities that the task is impossibly large almost no matter what the budget.

Fortunately, there are ways to handle this problem. One is "classification," which requires a new model for each identifiable class of relatively homogeneous commodities. The other is to specify as variables those attributes which identify the commodity within the model. In this case, commodity attributes would be used. Even though the number of variables required to do this is not trivial, it enables a single model to be constructed for a large group of commodities, which is very convenient. We may want to engage in limited classification of commodities as belonging to inherently large classes of shipments such as bulk liquid, dry bulk, refrigerated, general freight, etc., but the principal point is that a model for each commodity is impractical if it is to be broadly used.

Disaggregate models are particularly amenable to such an abstract commodity formulation. This is due to the previously mentioned reduced data collection requirements and the flexibility of a disaggregate choice model such as logit that includes handling different choice sets among observations and the possibility of specifying variables as generic or as alternative specific.

Alternative Freight Demand Models

There are several ways in which an individual shipper's behavior can be modelled. It was noted in section 2 that freight demand is an outcome of several choices. Existing models represent these choices in sequential steps (i.e., production and consumption, distribution, and modal split). However, it was noted that these decisions are often determined jointly and are interdependent in a way that makes a specific sequence assumption arbitrary. A more realistic representation of shipper behavior is a model that determines simultaneously choices which are highly interdependent. The most obvious deficiencies of existing freight mode choice models is their lack of explicit consideration of shipment size. Rail FCL, rail LCL, TOFC, common carrier truck LTL, etc., are all distinct alternatives that are considered by a shipper. Notice how the choice of mode and shipment size are intertwined in each of these options. Thus, a simultaneous model for the choice of mode and shipment size is the first step in improving the existing capabilities of freight demand models. Models that include choices of destinations can also be developed using the suggested modelling methodology. This is simply done by defining an alternative as a combination of mode, shipment size, and destination, for example. The model is formulated using the joint attractiveness of the combination of choices.

5. MODEL FORMULATION

In this section a freight demand model is developed from the concepts presented in the previous sections. The purpose of the model is to predict the volume of commodity k moving from origin i to destination j by mode m in shipment size q.

The Unit of Observation

The basic unit of observation in freight transportation is the individual shipment. Each shipment of a given commodity is characterized by its size and mode of transport, as well as its origin and destination. Therefore, the critical question is: How often does a shipper located at i transport commodity k to market j using mode m and shipment size q? To address this question, we can formulate the set of alternatives available to an individual shipper for the transport of commodity k and from i to j in the following manner:

alternative 1 - No shipment of k from i to j
alternative 2 - One shipment per month of k from i to j by FCL rail
alternative 3 - One shipment per week of k from i to j by FTL truck
alternative 4 - One shipment per week of k from i to j by rail TOFC
alternative 5 - Two shipments per week of k from i to j by LTL truck
.
.
.
alternative n - x shipments per week of k from i to j by mode m and shipment size q

Each alternative is described by giving the transport level of service attributes of that mode shipment size combination mq, the attributes of commodity k, the market attributes at j and the shipper attributes at i. It should be noted that because the "no shipments to market j" option is always open to a shipper, data must be collected on all potential markets for commodity k regardless of whether actual shipments are ever made.

The Model

Once the set of alternatives has been defined, the next step is to formulate a model for the probability that a shipper will choose a given alternative. It is assumed that the attractiveness associated with each

alternative is a function of its attributes and it is further assumed that the conditional probability of choosing an alternative is a function of its attractiveness and those of other alternatives which are available. Therefore, the model can be written as follows:

$$p^k(f,q,m,j|i) = g_{fqmj}(T,C,M,S)$$

where $p^k(f,q,m,j|i)$ = probability of shipper i transporting f shipments per week of commodity k by mode m in shipment size q to destination j.[5]

T = transport level of service attributes for all alternative q,m,j combinations for given i and k

C = attributes of commodity k

M = market attributes of all alternative j

S = attributes of the shipper of commodity k located at i

$g_{fqmj}(\ )$ = the choice model expressed for the probability of alternative fqmj

In practice, the model could be a multinomial logit function with variables, where $u_{fqmj}(T,C,M,S)$ is the joint attractiveness for the fqmj combination in which T and M take on the values for the specific fqmj alternative and C and S have the same values for all the alternatives.

Once the probability of frequency, mode, shipment size and destination has been determined for a given shipper of commodity k located at i, only one other piece of information is required for the forecasting of freight flows. The missing factor is the weight associated with a shipment of commodity k by mode m and shipment size q. In theory there is a distribution of weights associated with each mq combination for the shipment of k. However, if we define the modes and shipment sizes in alternatives in very specific terms, then we can approximate each

distribution of weights by its mean value. Furthermore, we can use data from shipments of commodity k between any origin and destination to estimate the mean shipment weight for each mode/shipment size combination.

Given a model for the probability of choosing any frequency, mode shipment size, and destination combination and given the mean shipment weight for each mq, then the desired model of freight flows can be written as follows:

$$V^k_{ijqm} = \sum_{\substack{\text{all shippers} \\ \text{of } k \text{ at } i}} \sum_f P^k(f,q,m,j|i) \cdot f \cdot Q^k_{qm}$$

where V^k_{ijqm} = tons of commodity k shipped from i to j by mode m in shipment size q

Q^k_{qm} = mean weight of a shipment of commodity k in size q on mode m

f = frequency of shipment

It should be noted that the calculation of V^k_{ijqm} involves a summation over all shippers of commodity k who are located at i. In most cases this summation presents no serious problem because so few shippers are involved. However, in some cases approximation procedures would have to be used. An estimate of total freight of commodity k originating at i, V^k_i, for example, can be obtained by a summation of V^k_{ijqm} over j, q, and m.

A model of this form can be estimated economically once proper disaggregate data can be assembled. The collection of this data is entirely feasible. However, little effort has been directed in this area to data. The next section addresses the sources of data that

are currently available and the following section describes a strategy for collecting the data which is still missing.

6. SOURCES OF INTERCITY COMMODITY FLOW DATA

The most widely available source of information on intercity commodity flows is the U.S. Census of Transportation.[6] This survey was first conducted in 1963, and it has been repeated in an expanded format in 1967 and 1972. The census contains an immense amount of information on the volume of freight traffic by commodity, by mode, by origin, by destination, by shipment size, and by length of haul. But, this set of data is of limited use for constructing disaggregate freight demand models of the type described here. The most obvious problem is that the Census of Transportation contains no level of service information. In practice, it is very difficult to use tariff books and carrier's schedules to deduce level of service information given only the freight flow data contained in the published volumes of the census. However, the Bureau of the Census has already made a pilot study to investigate the feasibility of collecting level of service data for publication in the Census of Transportation. But, even if level of service data is included in the next census, this data would be available no earlier than 1978.

The second problem with using the published volumes of the Census of Transportation for developing disaggregate demand models is that the data on freight flows is too aggregated. In order to protect the identity of individual shippers, the Census Bureau avoids stratifying the flow data by more than two or three variables at once. This makes it impossible to determine the shipper attributes described in section 2. Furthermore, the aggregation of shipments makes it very difficult to

deduce the alternative methods of transport which were not used. If the alternatives cannot be identified, it is impossible to model the demand for freight transportation using qualitative choice models of the type described in this paper.

The Census Bureau has released a computer tape based on 1967 Census of Transportation data cross-stratified by commodity, production area-market area pair, mode, and shipment size. The production areas and market areas are composed of small groups of SMSAs. A commodity cross-stratification is done at the 2-, 3-, 4-, and 5-digit commodity code levels. The mode breakdown includes rail, truck, air, and barge. No tariff travel time or other level of service data is included.

Although the census tape is very useful for other purposes, it falls short of being an ideal data set for disaggregate demand modelling. In the first place, many flows were not reported at the 5-digit commodity code level because less than five shippers were represented in the sample collected by the Census Bureau. Second, even the slight aggregation of shipments in the reported flows makes it difficult to identify transport alternatives and deduce shipper attributes. Nevertheless, this tape represents the most detailed freight flow information in general circulation. And, it is possible (with some difficulty) to gather some level of service attributes, as well as commodity and market attributes corresponding to flows reported on the census tape.

Some market attributes and general shipper attributes can be found in the U.S. Census of Manufacturing. For the large SMSAs in each state, the Census of Manufacturing contains production-related statistics for individual commodities at the 2-, 3-, and 4-digit code levels. These

statistics include the number of firms manufacturing the commodity, the total employment, wages paid, cost of materials used, value added, value of output, volume of output, and capital investment.

The origin, destination, mode, shipment size, and 5-digit commodity code description of a freight flow from the census tape can be used in conjunction with tariff books to estimate the cost of shipping by rail and truck. If the tariffs are not available, an alternative procedure is to estimate rates using a regression model.

Rail rates can be estimated from the data published in the DOT Rail Carload Waybill Statistics. This report gives the revenues per ton-mile for various shipment sizes and lengths of haul, for different commodities. The data is stratified at the 2-, 3-, 4-, and 5-digit commodity code level. However, the data is presented in state-to-state tables rather than on a city-to-city basis. Therefore, the rail rate must be estimated by comparing the exact length of haul with the average length of haul for the appropriate commodity and state, and adjusting the given revenue accordingly.

The truck statistics published by the ICC do not include enough specific information on revenues to allow the estimation of truck rates for the carriage of specific commodities between specific cities. However, truck rates can be roughly estimated from the data presented by Morton (1971) and his truck waybill study. These data consist of the average truckload rate for each of 174 weight and mileage blocks. If the length of haul is known and the shipment size can be estimated, then Morton's data can be used to estimate the rate.

Travel time and time reliability data for specific origin-destination pairs is very difficult to find. Rail travel time can be estimated from

train schedules, but travel time reliability cannot be estimated using this approach. There is no generally available information on truck travel times and time reliability for hauls between specific cities.

The scarcity of data is responsible in large part for the absence of many important variables from the freight demand models which have been published to date. However, several groups of researchers have partially overcome this constraint by collecting data for use in their freight demand studies.

Antle and Haynes (1971) collected a small disaggregate data set (eighty-seven observations) on barge, truck, rail, and combined barge-rail movements in the Upper Ohio River Valley. They concentrated their survey on firms receiving shipments of coal, coke, chemicals, and petroleum. From each receiver interviewed, they collected data on the shipment of one commodity by one mode, from one shipper. The characteristics included in the data are:

1. Annual tonnage of the commodity shipped
2. Distance
3. Average travel time
4. Average shipment size
5. Rate
6. Alternative rate
7. Handling cost

Unfortunately, they did not collect data on the average travel time, shipment size, and handling cost on the alternate mode.

The Army Corps of Engineers--South West Division has conducted a study similar to the Antle and Haynes study. They collected disaggregate

data on 195 barge, rail, and truck shipments of a variety of goods in the Arkansas River Valley. This data set includes the same variables as the Antle and Haynes data except that alternative rates for the nonchosen mode are not included.

Brian Kullman (1973) built a data set around the commodity flow information contained on the Census of Transportation computer tapes. He used the ICC Rail Carload Waybill Statistics to develop regression equations that can be used to estimate rail rates for certain commodities as a function of distance. He also used truck rate data from Morton (1971) to calibrate regression equations that can be used to estimate truck rates as a function of distance. Kullman obtained average rail transit times for city pairs in the northeast from Penn Central records. He used the same records to calculate the average rail time reliability for city pairs. He was unable to obtain similar information for trucks and so he estimated the truck transit times based on an assumed average daily mileage. The truck transit time reliability was assumed to be unity (i.e., perfect reliability). Kullman used Census of Manufacturing data to estimate the value of commodities by dividing the value of total output by the volume of output for selected commodity groups. Kullman's attempts to use this data to calibrate aggregate freight mode choice models, as described in section 3, met with little success due primarily to the measurement errors and the aggregate nature of this data set.

Hartwig and Linton (1974) collected 1213 freight waybills for full-load shipments of a particular consumer durable by rail and truck. From the bills, they determined the distance shipped, travel time, cost, shipment size, and value of the commodity being shipped. The relatively good empirical results reported by these researchers demonstrate the usefulness of accurate disaggregate data.

The Chicago Area Transportation Study has conducted a survey of firms which only have truck service available. This data set contains very little level of service information. However, the shipment attributes and shipper attributes that were collected are representative of the kind of data and level of detail that should be included in future data-gathering projects.

The list of shipper attributes collected includes the following:

1. Floorspace variables
   a. Office floorspace
   b. Manufacturing floorspace
   c. Total storage space
   d. Total plant floorspace
2. Employment variables
   a. Managerial employees
   b. Sales personnel
   c. Skilled employees--manufacturing
   d. Unskilled employees--manufacturing
   e. Service employees
   f. Other employees
   g. Total employees
3. Dummy variables
   a. Adequacy of small letter storage space
   b. Private truck availability
   c. Seasonal fluctuation of outbound shipments
   d. Seasonal fluctuation of inbound shipments

The data collected on individual shipments include the following:

1. Origin and Destination
2. For each commodity in the shipment:
   a. Volume
   b. Weight
   c. Packaging
   d. Handling
   e. Value
3. For the total shipment:
   a. Volume

b. Weight
c. Value
d. Transportation cost

The list of shipper attributes and shipment attributes covered in this survey could be expanded somewhat, but even so, this survey is much more detailed and comprehensive than most.

None of the existing disaggregate data sets contain enough information to allow the calibration of a complete disaggregate freight demand model. Most are based on a small number of commodities. Most of these data sets contain very few shipper attributes, market attributes, or transport level of service attributes. At best, most existing disaggregate data can be used to calibrate commodity specific models dealing with only the mode and shipment size decisions.

7. DATA REQUIREMENTS

Basically, four kinds of data are required for the estimation of a disaggregate freight demand model. The first kind of data consists of transport level of service attributes. Level of service data describe the cost and quality of service of each of the alternative market-mode-shipment size combinations. Market attributes are required for each potential destination from a given origin and for a given commodity type. The other data are commodity attributes and shipper attributes. These latter two kinds of variables are needed to determine the emphasis that a particular individual shipper will place on each of the level of service and market attributes when making the transport (frequency, mode, shipment size, and market) decision. Market attributes vary across origin-destination pairs and explain the frequency of shipment and market choice of a given commodity as well as act together with the commodity and shipper

attributes in explaining the values placed by a shipper on the level of service attributes.

Constraints on Data Collection

The discussion of the preceding section leads to several recommendations for future data collection efforts. The most important is that data must be gathered on the transport level of service attributes which would be experienced by a particular shipment: both for the market-mode-shipment size combination actually selected and for the alternatives which were not chosen. Since the shipper is choosing between alternatives, the full list of attributes for all alternatives, whether or not they have been chosen, is necessary. Second, the collection of freight flow data should be undertaken in conjunction with shipper interviews in order to collect simultaneously commodity and shipper attributes. Third, data should be collected on a full range of different commodities, modes, and shipment sizes.

Shipper Interviews

The level of service data is the most difficult to collect. The shipper may have good information on the level of service for the market-mode-shipment size option which he most often uses. However, his knowledge of the level of service associated with the other options may be considerably less accurate, depending on the frequency with which these other (nonchosen) options are used. The shipper can probably define most accurately the set of truly competitive options. Certainly shippers can help identify the set of potential markets, but shipper interviews alone are not enough to determine accurately the attributes of each of the available options.

different commodities, modes, and shipment sizes.

Shipper Interviews

The level of service data is the most difficult to collect. The shipper may have good information on the level of service for the market-mode-shipment size option which he most often uses. However, his knowledge of the level of service associated with the other options may be considerably less accurate, depending on the frequency with which these other (nonchosen) options are used. The shipper can probably define most accurately the set of truly competitive options. Certainly shippers can help identify the set of potential markets, but, shipper interviews alone are not enough to determine accurately the attributes of each of the available options.

Note also that there is a difference between the shipper who is acting as the consignee (receiver) and the one who is functioning as the consignor of a particular shipment. The consignee will probably be the better of the two to interview because he is more likely to have travel times and completed shipping documents. However, for commodities for which the shipment decision is made at the origin the consignor must also be interviewed.

Analysis of Waybills

Waybills or bills of lading will probably be the primary source of level of service data for the mode actually chosen. These documents indicate the identity of the commodity shipped, the shipment size, and mode chosen. They also include the tariff. They may also indicate the time spent in transit on the primary mode of carriage.

However, waybills do not include information on the costs of packaging, loading, and unloading. Nor do they include the cost of access to and from the primary carrier's facilities. Waybills cannot be used as the only source of level of service attributes, however, because they do not describe all of the options available for the shipment of each commodity for each origin-destination pair, for all commodities and O-D pairs being studied. This is best determined by individual investigation of the shipping possibilities in conjunction with the shipper interview.

The other prinicipal purpose of gathering waybills is to find the frequency at which a mode, shipment size, and market combination is used by a shipper for a given commodity. To determine frequency, waybills should be collected for all shipments from a given shipper for a given span of time, such as one week. This type of waybill survey should be repeated for each shipper in the data set. Alternative frequency information can be obtained at a reduced level of accuracy by direct questioning during the shipper interviews.

Carrier Interviews

Interviews with carriers may represent the most straightforward and accurate method of gathering level of service data for those modes not chosen. The carriers can provide the expected travel time between their facilities, as well as estimates of the travel time reliability. However, much of this data can also be estimated from the consignee's records if any shipments were actually made. The carriers tariff schedule can be used to find the cost of using

various mode/shipment size combinations for the shipment of a given commodity. However, there are several factors which complicate the calculation of shipping costs from carrier data. One problem is that many shipments involve hauls in several tariff regions. Most carriers have tariffs only for their own regions. This means that a large tariff library such as the ICC Library in Washington, D.C. may have to be used. The problem is further complicated by the existence of many "exception: rates for specific commodities in particular tariff regions. All in all, the determination of shipping costs for each of the options will be a formidable task.

Another aspect of the level of service data which might be available from carriers is information on loss and damage claims. This information is tabulated for insurance purposes, but carriers may be reluctant to release the figures. They may also have relevant information on commodity shipping attributes such as density, value per pound, shelf life, etc.

Data Summary

In summary, the data on transport level of service attributes could be collected in three steps. First, interview shippers to determine a reasonable set of market, mode, and shipment size alternatives and, at the same time, gather data on packaging, handling, and access costs, etc. Second, collect from them copies of all waybills for a specific time period to find the frequency of shipment and which alternative is actually chosen. Third, consult carriers to assist in determining travel time, time reliability, and shipment costs on the alternatives which were not chosen.

The amount of data to be collected must be assessed more carefully than we can do here. By current data collection standards (the 1% Waybill Study, the FHWA Truck Study, or the Census of Transportation Commodity Shipment Survey) this effort will be very small, perhaps as few as 2500 to 5000 data points. Once embodied in the model, however, the data will be transferable from one region to another, from one time to another and, hopefully, from one commodity to another. It will therefore be extremely cost effective.

8. CONCLUSIONS AND RECOMMENDATIONS

The need for a policy sensitive model of freight demand has risen dramatically over the past few years. Questions of railroad bankruptcy, line abandonment, energy conservation, deregulation, rising food prices, improvement of air quality, land use controls, offshore ports, channel dredging and multimodal coordination all impact directly on freight transportation as it has been traditionally performed. Methods for determining the nature of these impacts and for assisting the policy-maker to minimize negative impacts have not existed in the freight area. Likewise, opportunities to exploit new technologies or revised institutional structures to improve the efficiency and performance of the existing system have not been adequately addressed. Part of this failure to explore new opportunities arises because adequate forecasting tools have not existed.

Existing freight demand models are not sensititve to a majority of the policies which are under consideration. They employ a very restrictive set of variables and predict very limited information. The result is a lack of credibility for freight forecasting generally, by both public and private decision-makers.

There are two major reasons why this condition has existed. First, an adequate framework for explaining the phenomena of freight transport and model forms for predicting it did not exist. Second, and partly as the result of the lack of theory, data on freight transport has been notably deficient. Data sets that are currently available could, at best, be used for only a preliminary shipment size, mode choice model for a restricted set of commodities and modes.

New data collection -- on a small scale -- but with more complete information is required. This paper describes an approach to freight demand modelling in sufficient detail that the data requirements can be specified precisely. A first look at the size and difficulty of organizing and conducting the data collection suggests that it can be done economically and that the results will more than justify the means.

The modelling methodology proposed here makes the development of an operational set of policy-sensitive models feasible once the proper data has been collected. In fact, the feasibility of collecting the data for use in estimating the model depends rather importantly on several features of the models which greatly increase their scope of application. These are:

1. Efficiency - use of disaggregate techniques allows the full range of variability to be attributed to the behavioral characteristics of the shipper, which means that much less data is required than in aggregate, zonally based approaches
2. Transferability - models built from small samples of data collected in one area of region of the country can be used in other areas, as well as transferred over time
3. Aggregation - once the behavior of individual shippers is known, models can be aggregated to predict freight demand at more aggregate levels

4. Commodity abstractions - one model (or, at worst, a small set of them) is built involving commodity attributes, avoiding having to gather data, specify, estimate, and test a model for each commodity for which predictictions are desired -- an impossible task

The model described here represents a major advance over the present state of the art. As such, it is necessary to reevaluate the usefulness of freight demand models in the study of policy alternatives for the national freight system. If the conclusion is positive, then data collection to support the development of such a model can begin.

NOTES

1. As Noted, there is a feedback relationship between the actions of users and suppliers of freight transport. This discussion, however, takes the point of view of an individual user who is assumed to face given supply characteristics.

2. The Bureau of Economic Analysis (Department of Commerce), has divided the United States into 173 BEA Regions. These are identical to the 173 OBE old Regions, formerly the Office of Business Economics (Department of Commerce).

3. Quandt, R. E. and E. H. Young, "Cross-Sectional Travel Demand Models: Estimation and Tests," Journal of Regional Science, Vol. 9, No. 2, 1969.

4. For a more detailed description of this modelling methodology see CRA (1972) and Ben-Akiva (1973).

5. For commodities for which the shipment decision is made at destination i, j should be interchanged such that the model predicts the probability of choice of origin for a given destination.

6. Department of Commerce, Bureau of the Census, 1967 Census of Transportation.

REFERENCES

Antle, L. G. and Haynes, R. W., "An Application of Discriminant Analysis to the Division of Traffic Between Transport Modes. "Springfield, Virginia: Institute for Water Resources, Report 71-2, NTIS, Department of Commerce, 1971.

Arthur D. Little, Inc., "Domestic Waterborne Shipping Market Analysis - Appendix C - Modal Share Estimates." Prepared for Maritime Administration, U.S. Department of Commerce, 1974.

Hartwig, J. and Linton, W. "Disaggregate Mode Choice Models of Intercity Freight Movement." Master's thesis, Northwestern University, 1974.

Institute of Water Resources, "Discriminate Analysis Applied to Commodity Shipments in the Arkansas River Area.." Prepared by Army Engineer Southwestern Division, Dallas, Texas; Report IWR-RR-74-R2, 1974.

Kullman, B. "A Model of Rail/Truck Competition in the Intercity Freight Market." Ph.D. dissertation, Massachusetts Institute of Technology, 1973.

Mathamatica, "Studies on the Demand for Freight Transportation." PB 176 479, vol. 1. Mathamatica, 1967: Princeton, N.J.

Morton, A. "Competition in the Intercity Freight Market." Washington, D.C. Office of Systems Analysis, Department of Transportation, 1971.

Morton, A. "Truck-Rail Competition for Traffic in Manufactures." Transportation Research Forum Proceedings, 1971.

Perle, E. "The Demand for Transportation: Regional and Commodity Studies in the United States." University of Chicago Press, 1964.

Reebie Associates. "Truck/Rail Intermodal Operations - An Option for the Future?" Prepared for Department of Transportation, 1974.

Sloss, J. "The Demand for Intercity Motor Freight Transport: A Macroeconomic Analysis." The Journal of Business, 1971.

U.S. Department of Commerce, Bureau of the Census, Census of Manufacturing 1967, Industry Statistics Vols. 1,2, and 3. Washington, D.C.: U.S. Government Printing Office, 1971.

U.S. Department of Commerce, Bureau of the Census, Census of Transportation 1967, Vols. 1,2,3,4. Washington, D.C.: U.S. Government Printing Office, 1971.

U.S. Department of Transportation, Carload Waybill Statistics, 1969. Washington, D.C.: Office of Systems Analysis and Information, 1971.

Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9

NEEDS AND PRIORITIES IN RESEARCH ON RAIL SERVICE RELIABILITY

Joseph M. Sussman*

I INTRODUCTION

The rail industry is a complex one. Additionally, when one considers the interactions of this complex industry with the external world of competition, regulation, and general economic trends, the situation becomes difficult, if not impossible, to analyze. One attempt to represent the various relationships is shown in figure 1.[1]

There are a number of key interactions embodied in this representation. Among the most important are:

1. The relationship between the quality of rail service offered by the industry to its shippers and the volume of traffic captured by the railroads (with the obvious implications on rail revenues and profitability)
2. The relationship between operating policies, service levels, and operating costs (with, once again, the impact on industry profitability)
3. The relationship between the strength of, and service supplied by, the competitive elements (rail, barge, pipeline) and the volume and mix of the traffic captured by the rail industry
4. The relationship between the profitability of the industry, investment levels, and the quality and quantity of rail facilities--a relationship that involves the ability of the industry to replenish itself as its capital is used up

*Professor, M.I.T.

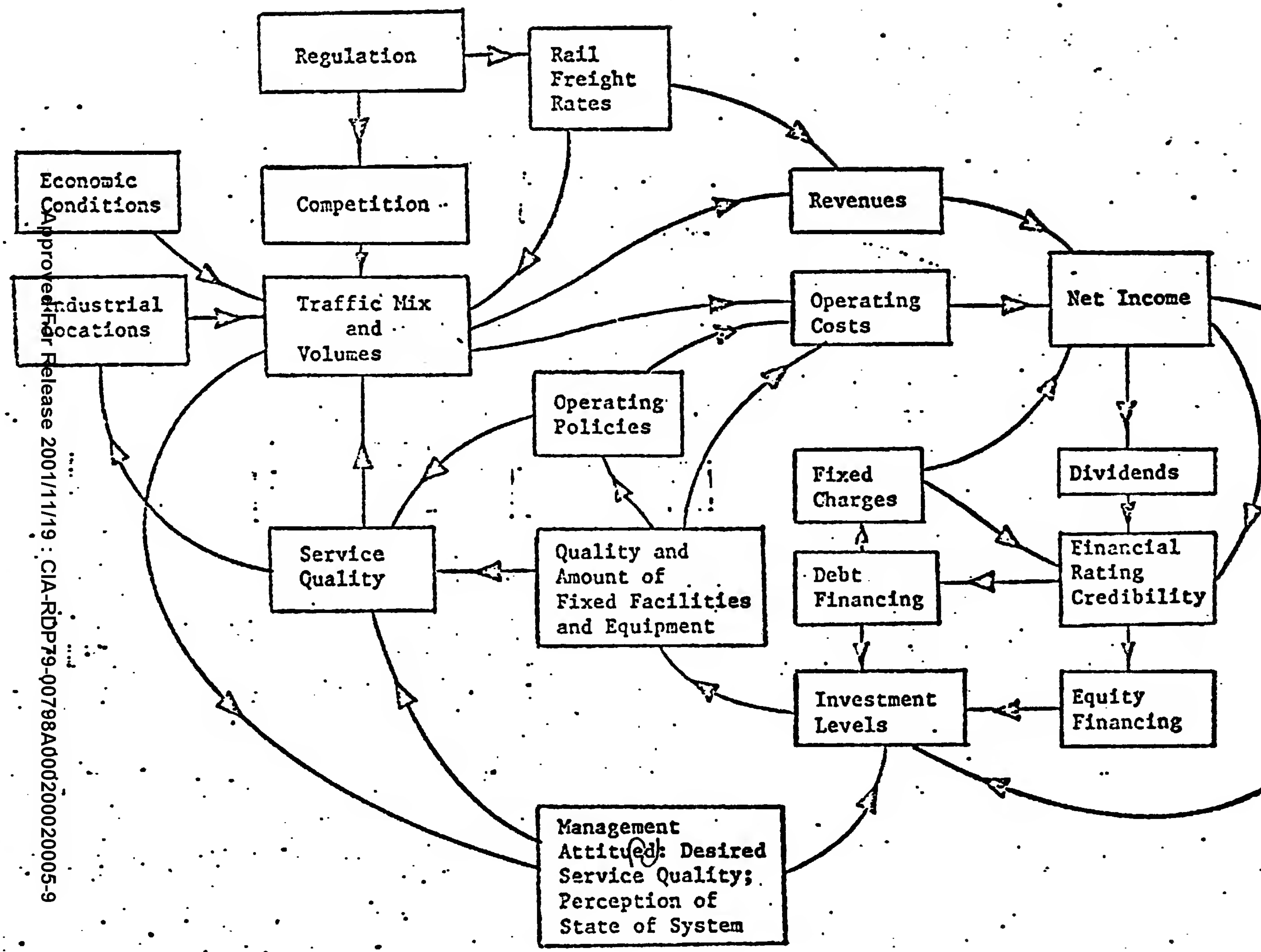


FIGURE 1 RAIL INDUSTRY SYSTEM REPRESENTATION

This paper deals with research needs and priorities in the field of rail service reliability. However, it should be clear from the previous discussion that any approach to this topic must take into account the relationship between service reliability and the many factors which impinge upon it and upon which it impinges. In fact, as the reader will see, many of the major research areas in the future deal, not simply with methods of improving service reliability but rather with the relation between service reliability and other factors (e.g., traffic volumes attracted by the rail industry as a function of service reliability).

II SERVICE MEASUREMENT

While the focus of this paper is service reliability, the author would be remiss if he failed to point out that service reliability is but one of the many level-of-service parameters that the shipper may use in judging the viability of the railroad's transportation product. Some others in addition to service reliability are average transit time (between particular origin-destination pairs), empty car availability, loss and damage, and freight rates.

The rational shipper judges the railroads by the total economic impact on his business or using this mode. Thus, he is concerned with service, starting from the time he orders empties until his goods are delivered to the consignee. He is not interested in the car cycle (time between consecutive loadings of a freight car) or in on-time performance of a particular symbol train. Rather, he is interested in how long it takes him to get an empty car spotted at his sidings (and whether it is a usable car) and in how the railroad handles his car between origin and destination. To some extent, the industry has focussed on the former, internal measures of rail service and not on the latter, shipper-oriented measures.

As some examples of origin-destination performance in the rail industry, consider figure 2.[2] Those histograms (taken from actual rail data) show rail performance between six origin-destination pairs. Clearly, all cars travelling between a particular origin and destination do not exhibit the same performance. This variation in performance is measured by the so-called 3-day %, the maximum number of cars arriving in a three-day window. "3-day %" is one measure of shipper-perceived service reliability. Martland[3] describes several other measures and discusses their usefulness.

Important to note is that mean transit time can be constructed as a measure of rail performance separate from service reliability. The hypothesis is that the shipper perceives them as distinct entities. However, both the mean transit time and the 3-day % can be translated into costs for the shipper. For example, figure 3 shows (for a hypothetical situation how improvements in 3-day % (i.e., improvement in reliability) reduces the reorder point in a typical inventory system, which will typically reduce the shippers' logistics costs. Figure 4 shows that, at a given reliability level (3-day % = 80%), reduction of mean transit time below a certain threshold (in this case, three days) does little for the shipper in terms of reduced logistics costs.

III SOME RESEARCH RESULTS

M.I.T.[4] has carried out a research study of rail service reliability and the conditions under which unreliability is likely to occur. This section briefly notes some of these results.

First, there is a clear relation between average transit time and transit time reliability. Specifically, figure 5 shows that trips with

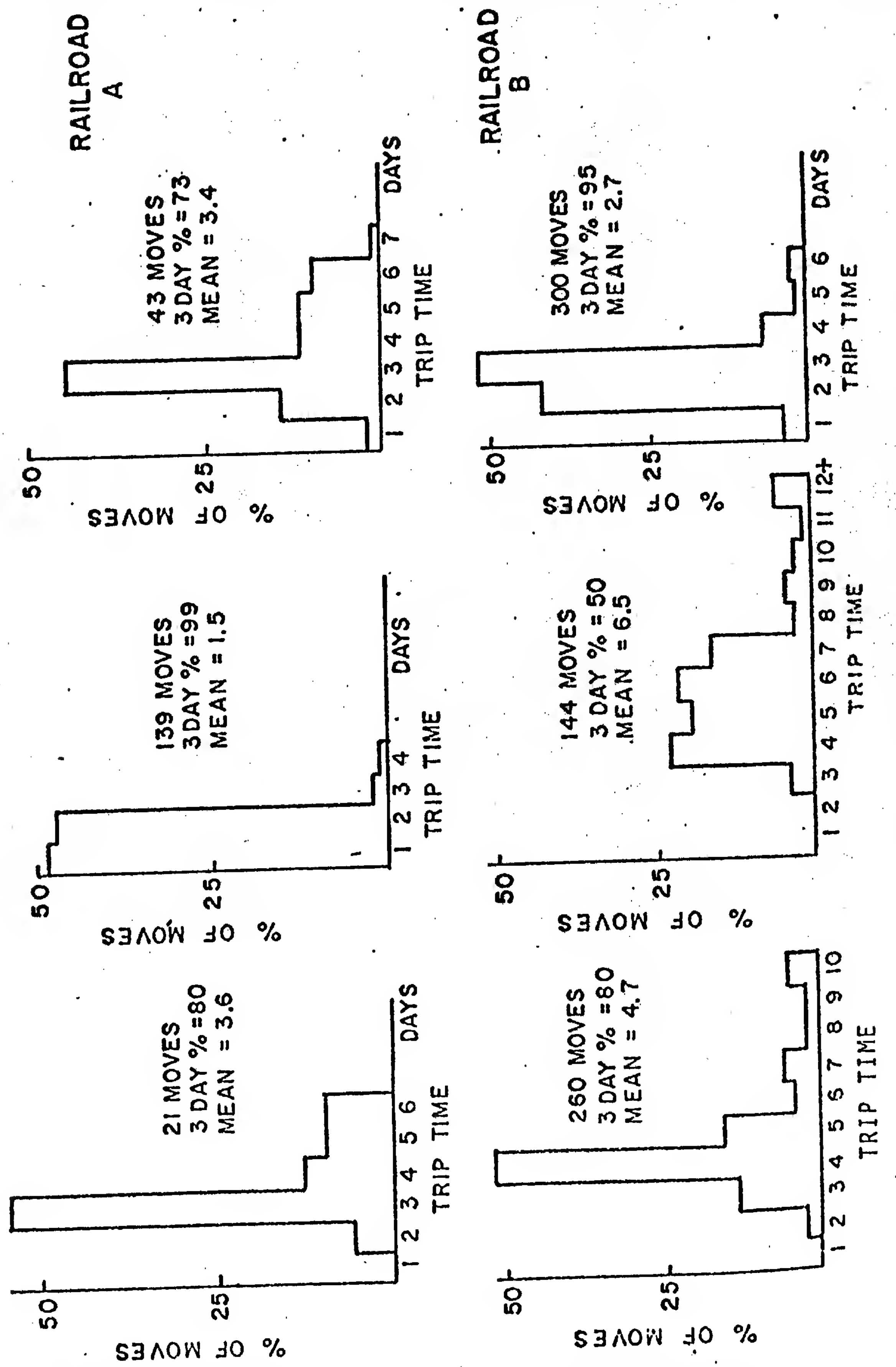


FIGURE 2 SAMPLE O-D DISTRIBUTIONS

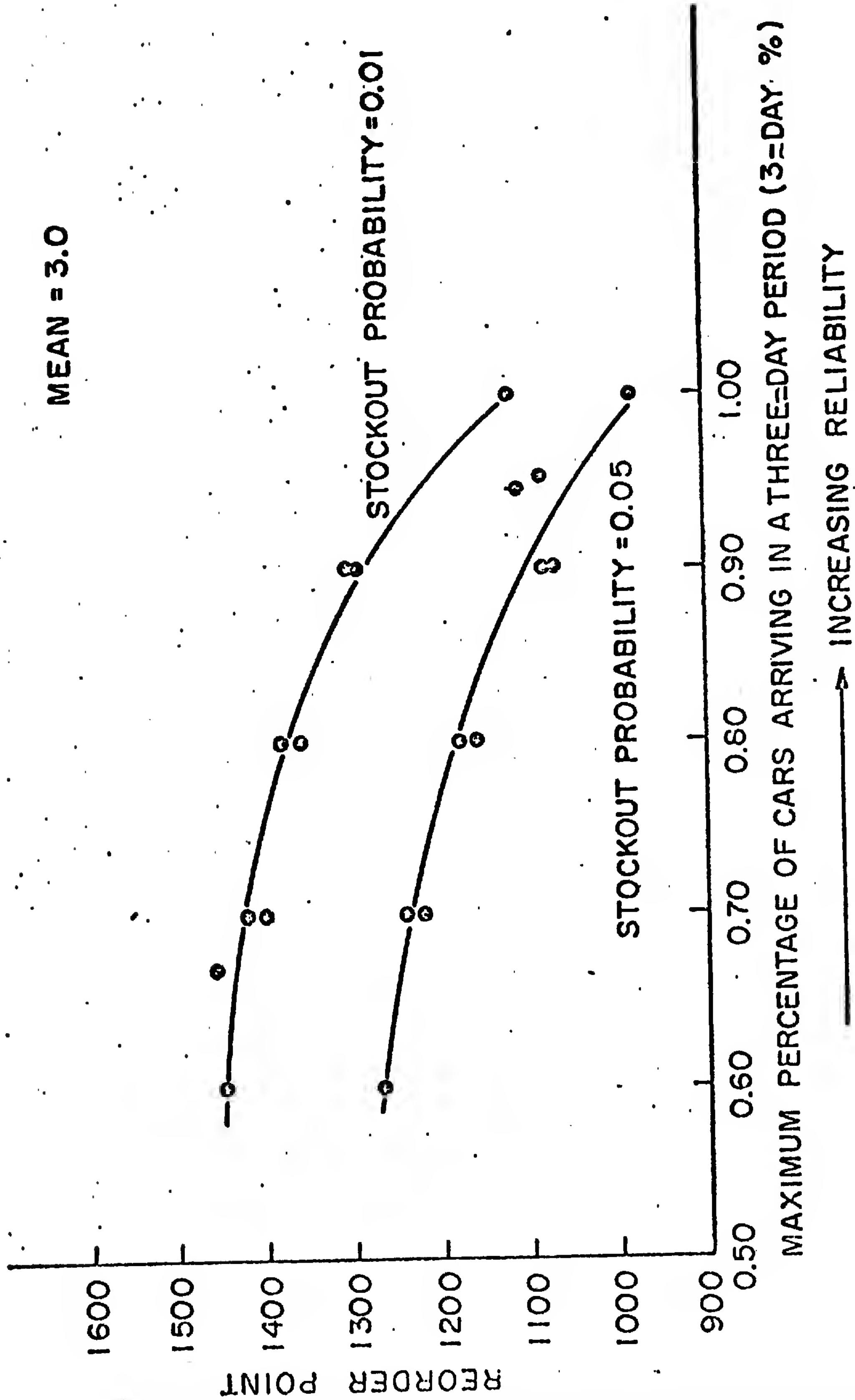


FIGURE 3 EFFECT OF CHANGING 3-DAY %

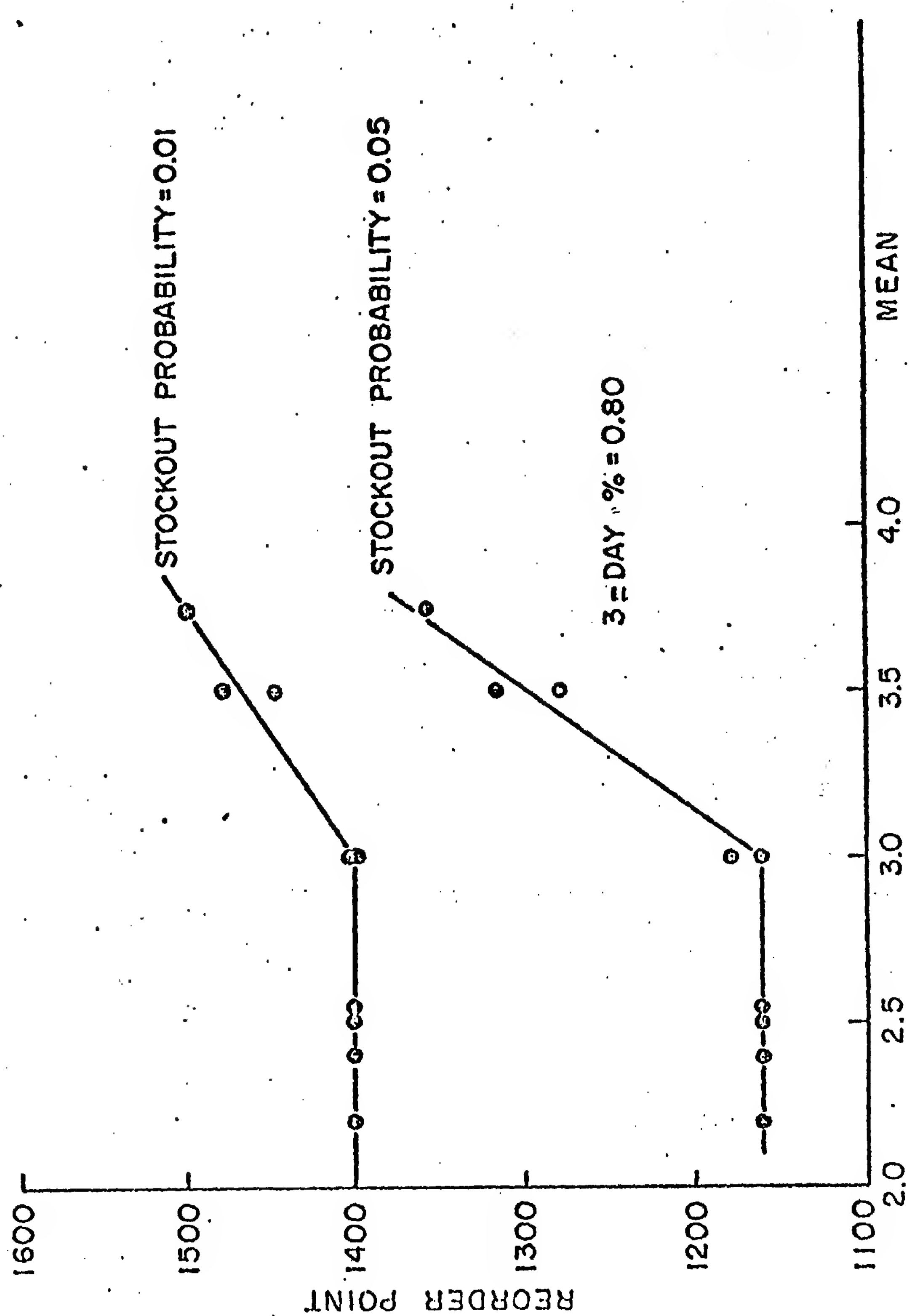


FIGURE 4 EFFECT OF MEAN TRIP TIME

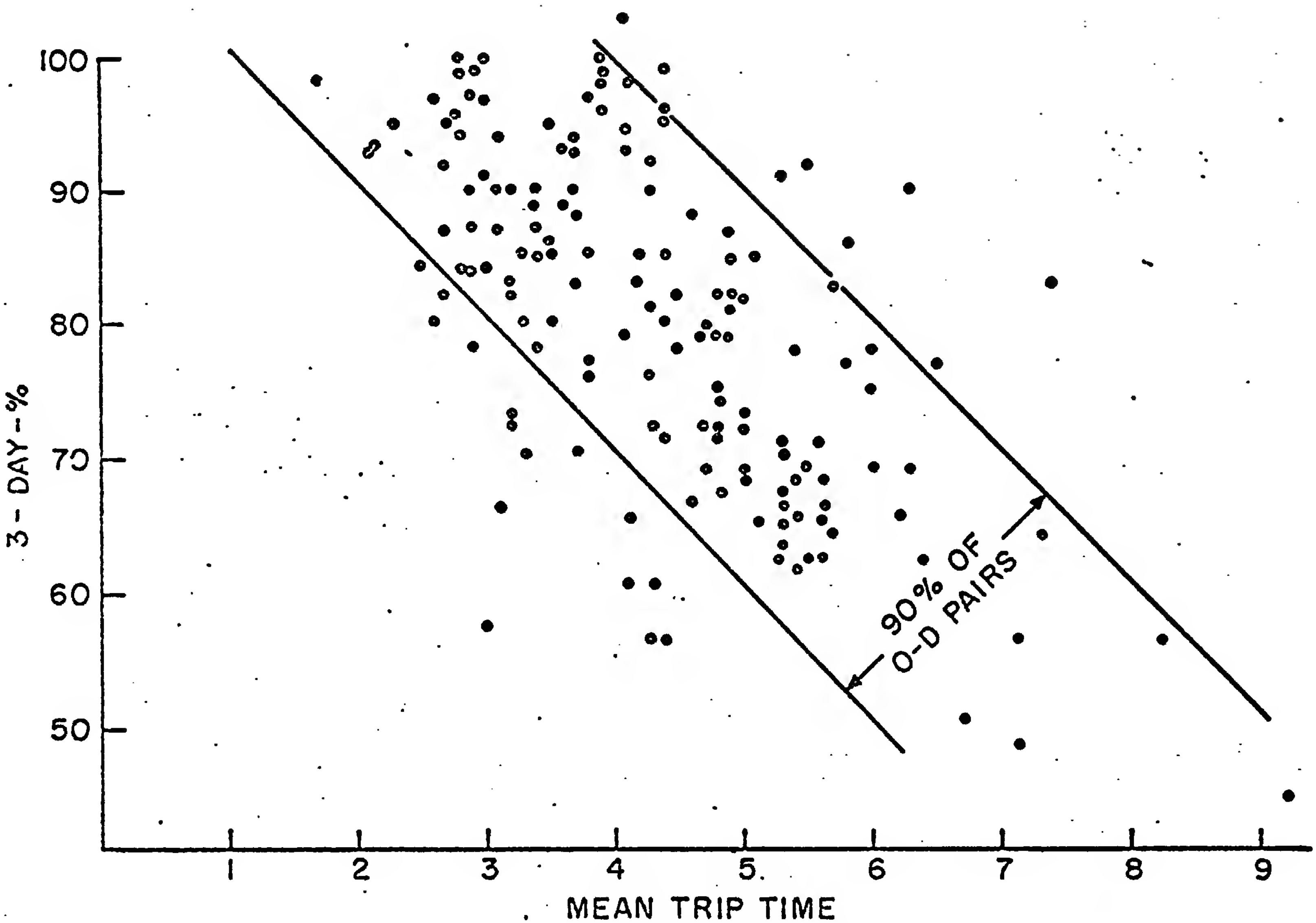


FIGURE 5 THE RELATIONSHIP BETWEEN RELIABILITY AND THE MEAN TRIP TIME (EACH DOT REPRESENTS ONE O-D PAIR)

high mean transit times tend to be the most unreliable. Perhaps more meaningful from a rail operations perspective, it appears that the number of terminals at which a car is handled is the major determinant of performance. For example, consider figure 6. σ^2_{LH} [5] is a measure of variation in transit time. Those moves with zero intermediate yardings exhibit slight variation and hence tend to be reliable. Those with two intermediate yards are substantially less reliable. On the basis that moves with more yardings are the long distance moves in a system, one might argue that distance is the cause of the unreliability. Some evidence that these reliability differences are caused by the number of intermediate yards rather than the distance is shown in figure 7.

That this relationship between terminals and origin-destination service exists is not surprising when one considers rail operations. It is the nature of railroad operations that a car encounters numerous opportunities for delay as it moves from its origin to its final destination. At each yard, cars moving to common, intermediate, or final destinations are consolidated into "blocks," placed in a train consisting of one or more blocks, and handled together to another yard which may be twenty or more than a thousand miles distant. Whenever a car is set off from a train or the train reaches its destination, the car is reswitched and consolidated with other traffic into a new block and a new train. This procedure is repeated until the car reaches its final destination.

This process of switching and consolidation necessarily results in longer transit times than would be required for direct movement (such as by unit train). Equally as important, this process is unreliable. That is, each time a car is switched, the potential for a missed connection at that yard exists.

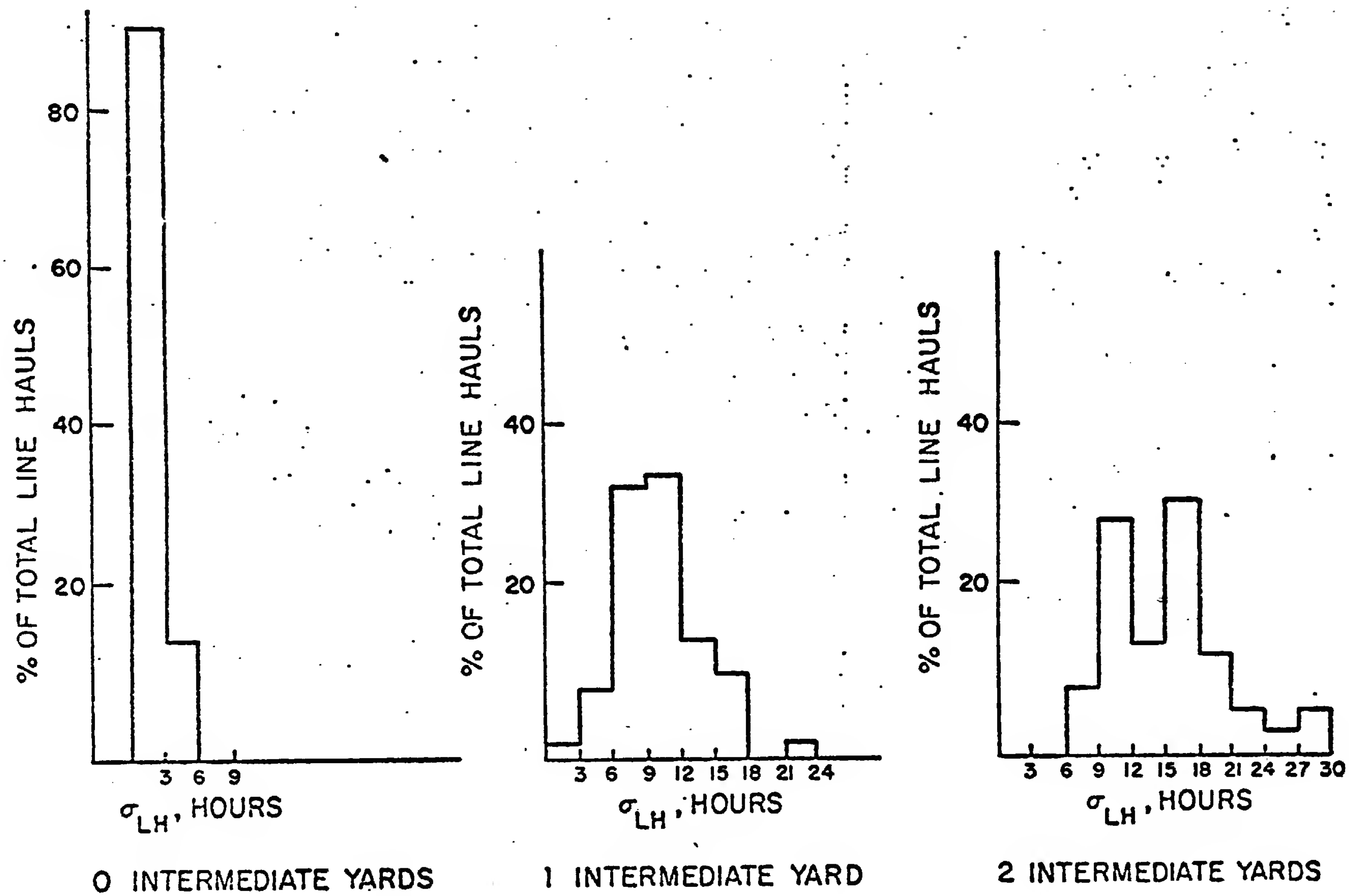


FIGURE 6 INTERMEDIATE YARDS AND TOTAL LINE HAUL RELIABILITY

Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9

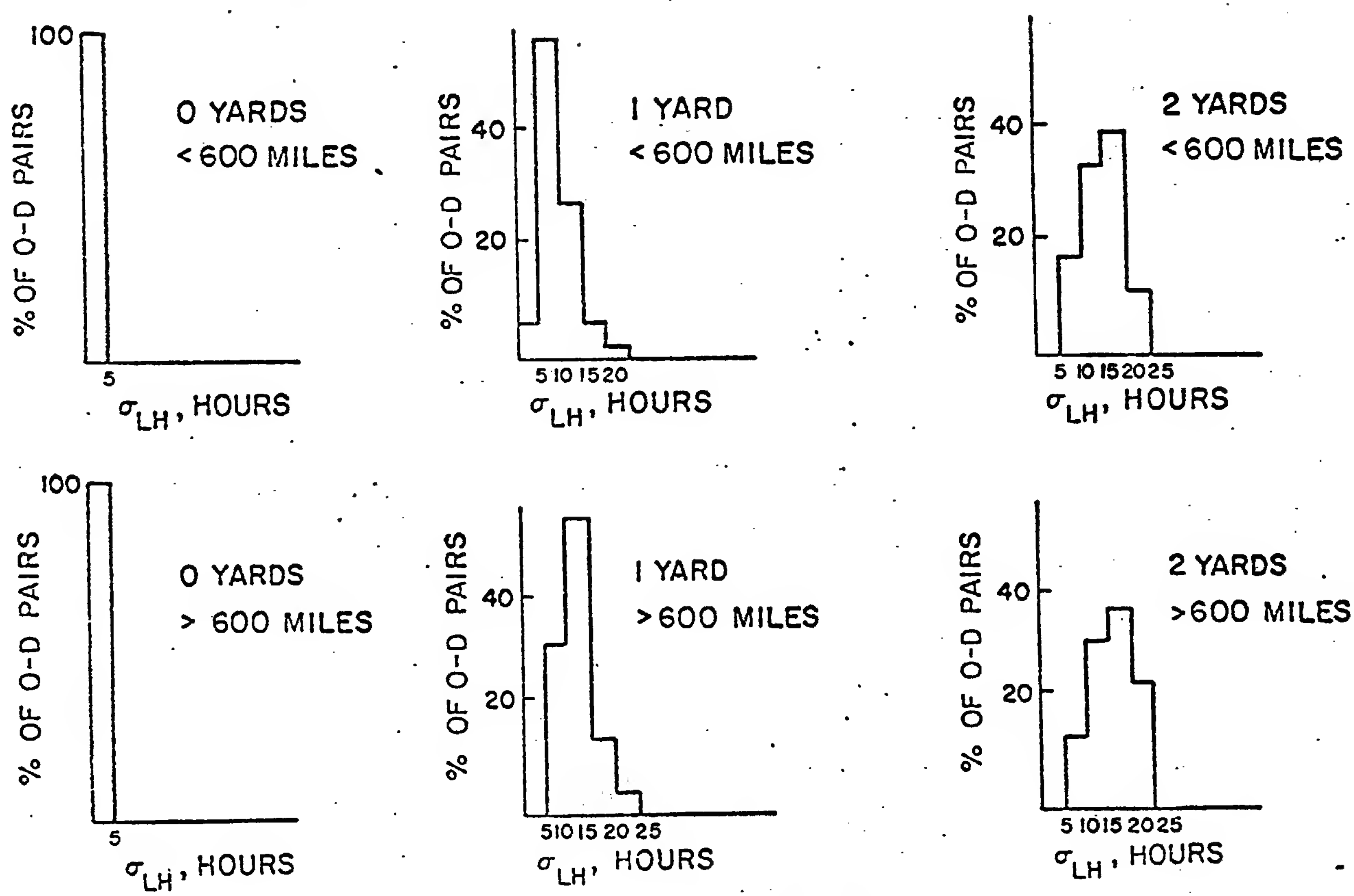


FIGURE 7 THE PREDOMINANCE OF THE NUMBER OF INTERMEDIATE YARDS OVER DISTANCE AS A FACTOR AFFECTING TOTAL LINE HAUL RELIABILITY

Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9

Missed connections are critical in that they often lead to car delays in the order of twelve-twenty-four hours (the time until the next appropriate outbound train), large variations in transit time and, hence, unreliable performance. Table 1 shows the magnitude of transit time delay and transit time variance as a function of the probability of a missed connection at a yard. These probabilities are quite realistic in the light of the various analyses of railroad operating data.[6]

TABLE 1

AVERAGE DELAY TIME AND STANDARD DEVIATION OF DELAY TIME AS A FUNCTION OF THE PROBABILITY OF A MISSED CONNECTION

Probability of Missing Connection	Average Delay (Hours)	Standard Deviation of Transit Times (Hours
.1	2.4	7.2
.2	4.8	9.5
.3	7.2	11.0

Among the causes for missed connections are outbound train cancellations, train length/weight constraints, RIPS, NO-BILLS, and late arrival of an inbound car. This last cause has been shown to be an important one. Specifically, if the car arrives later than some "threshold" time, its connection with an outbound is often missed. Of course, the outbound train could be held for the car, allowing the connection to be made despite the lateness of the arrival. However, this may well lead to further problems. Specifically, Belovarac and Kneafsey[7] have shown that the primary cause of late arrivals at a yard is late departure from the preceding yard. Hence, holding trains to allow particular connections to be made may well lead to inbound lateness at succeeding yards and the possibilities of other missed

connections. Indeed, Folk[8] had examined "no-hold" versus "hold" policies with respect to overall network effects on car performance.

All in all, the question of yard performance and missed connections is a complex one, with the various components and operating policies of the rail network heavily interacting to affect performance.

These research results indicated that potential for improvement in service existed through operating strategies that would either avoid intermediate yardings of cars (i.e., run-through trains) or improve the probability of making connections when yardings were necessary (e.g., by optimizing inbound/outbound train connections or by limiting cancellations). A case study was performed in cooperation with the Southern Railway[9] and various experimental alternatives designed to improve service were implemented. These implementations did lead to service improvements in the Southern's system and have been integrated into the Southern's continuing operating plan.

IV RESEARCH NEEDS

With the previous as background, this paper goes on to discuss research needs of the rail industry in the area of service and service reliability.

The research needs that exist are subdivided into two major areas as follows:

1. The shipper's perspective
2. The supply of transportation service

IV-A THE SHIPPER'S PERSPECTIVE

The area can be further subdivided into the shipper's demand for transportation service and the shipper's operating response to rail service.

IV-A-1 DEMAND

From the viewpoint of the author, the major gap in rail service research (and, in fact, in freight service in general) is on the demand side of the question. Specifically, we have only the most preliminary and intuitive ideas about the relationships between level of service provided by the industry and the volume and traffic mix attracted by the industry. There is a great shortage of knowledge about what service the shipper community actually wants and needs and, more importantly, how the shipper will make his model choice decision as a function of the service provided. Making major investment decisions or operating decisions in the absence of demand models (i.e., models for understanding the volume complications of service decisions) seems quite difficult. Yet, this seems to be what the industry does (although specific instances of demand sensitivity or insensitivity have been documented).

Thus, a major research priority should be the development of a level-of-service and commodity-sensitive demand model for use as a planning tool by the rail industry in making investment decisions. Of potential major benefit in this endeavor may be newly developed, disaggregate, demand-modelling methodology. This methodology has the potential for overcoming what has long been the major roadblock in the development of such models; namely, the lack of an adequate data base.[10]

We should note that there is by no means complete agreement in the industry on what constitutes service and how service relates to demand. For example, in informal discussions, various rail officials have indicated that the shipper does not see himself as "buying" reliability from the industry and that modest changes in volume (if any) would occur were service reliability to be dramatically improved. By the same token, other

rail officials have indicated that service improvements can lead to substantial attraction of new business. In fact, neither contention has been supported by a comprehensive set of data, models, and analysis.

In addition to the obvious needs for data, other preconditions exist for success in this effort. Perhaps most important is the market's perception of what rail service is. In order to develop consistent and usable demand models, performance measures that are useful from the shipper's perception must be developed. Of course, these measures must exist (or be capturable) in the data base used in model calibration.

In summary, to quote from existing work:

> If a railroad does not know how service levels affect traffic volumes, then that railroad does not know what product it should produce.[11]

Such a modelling effort will allow the industry to better understand this issue.

IV-A-2 SHIPPER'S OPERATING RESPONSE TO RAIL SERVICE

The previous section focussed on the attraction (or lack thereof) of traffic as a function of service. A second shipper-related area is composed of the operating behavior of a shipper who is, in fact, using rail and how this behavior may be related to the quality of rail service.

Shipper behavior and procedures are an important component in the utilization of the rail industry's resources. For example, consider car detention. A not-insignificant portion of the car cycle is time spent with the car in shipper control.[12] An empty car may wait to be loaded for some time and a loaded car may be used as a rolling or stationary warehouse by the consignee. Often demurrage is not an adequate incentive for the shipper to load or unload the cars.

The working hypothesis here is that this shipper behavior is closely related to the service provided by the industry. For example, one reason advanced by the shipper community for "excess" car detention on loading is that poor car-availability leads to very conservative car-ordering policy. This, in turn, may lead to empty cars waiting for long time periods at the shipper's siding if they are, in fact, delivered in a timely fashion. Unrealiable delivery of empties or spotting of loads may also lead to "erratic" shipper behavior because the shipper has difficulty in scheduling his crews around the undependable service provided by the industry.

The above suggests that improvements in service (and service reliability in particular) may pay extra dividends in terms of an improvement in the car detention situation. That is, if the rail industry provides more reliable service, equipment utilization can be improved because of better shipper performance which can, in turn, lead (through better availability of equipment) to still better rail services.

Another related area is the regularization of shipper demands on the rail system. The irregularity of shipper demands causes poor allocation of resources as the rail system attempts to respond to peak demands for service (difficult to predict). This poor distribution of resources leads in turn to degradation of service on the network taken as a whole. Regularization of shipper demands can potentially lead to better service for all shippers. Again, a feedback phenomenon exists. Possibly, if service reliability can be improved, shippers can be enticed to regularize their inputs to the system, which will allow the rail industry to improve service still further.

The above are but two examples of the interaction of the shipper with the rail system. In all, there appears to be a great deal we do not know about shipper behavior patterns and how these relate to rail service. A research effort in this area can yield very meaningful results from the viewpoint of improved service and shipper satisfaction with rail service.

IV-B THE SUPPLY OF TRANSPORTATION SERVICE

The improvement of the transportation service provided by the rail industry can take place on three fronts. These are:

1. Rail Operations
2. Capital Expenditures on Rail Facilities
3. Institutional Changes

To put these needs in perspective, it's important to note that the previously described M.I.T. research, virtually all of which is supply-oriented, has merely scratched the surface of the field. This research has

1. Described some service measures
2. Isolated some of the root causes of rail service unreliability
3. Developed various models that can be useful in improving rail service
4. Demonstrated the usefulness of these models in <u>one</u> particular operating theatre

This research has had major limitations. It has focussed on the <u>operating theatre</u> only and, further, (in its implementation) has focussed on local and individual changes to operations and <u>not</u> on global changes to network operations.

In short, there is a great deal to do in research in this area. The remainder of this section goes on to describe these research needs on the supply side.

IV-B-1 RAIL OPERATIONS

It is clear that changes to rail operations can cause major impacts on rail service. What is less clear is how best to proceed and how much these changes are likely to cost (or save). Three areas are discussed, these being:

1. Impacts of Network Wide Operating Changes
2. Data Systems
3. Cost Models

Impacts of Network Operating Changes

The previous research has shown that local changes in rail operations (e.g., changing a single inbound/outbound train connection) can lead to the desired service effect. What needs to be better understood is the benefits (from a service perspective) of wholesale, network-wide changes to rail operations. Models for the prediction of the benefits of implementation of:

1. Blocking policy changes
2. Scheduling policy changes
3. Train length policy
4. Through trains

carried out on a network basis are badly needed. It should be noted that a wide variety of models that begin to address these issues are available. Network simulation models have been developed by many railroads and by the industry through the A.A.R.[13] These models allow an analyst to simulate rail operations at very fine levels of detail and can, in theory, be very useful in the examination of such issues as those noted above.

However, it is the author's impression that these models have not been particularly well utilized in address

questions. Why this is true (if indeed it is) is not clear. While these models have sometimes been criticized for being too expensive or difficult to use, it seems obvious that the costs of running the models will be trivial compared with the costs of implementing some operating change in the rail system. Given this situation, it is recommended that the micro-simulation approach to rail network modelling be reexamined afresh to ascertain why efforts along these lines have not been particularly fruitful to date in the rail industry (unlike the experience in other industries such as air transportation, manufacturing systems, etc.). If, in fact, conceptual problems with the microscopic approach can be isolated, further research on correcting these deficiencies is indicated.

In parallel with this reexamination of the traditional approach to rail systems modelling, some new approaches are appropriate as well. For example, less detailed network modelling approaches have proven fruitful overseas. In particular, ROUTESTRAT,[14] developed by British Railways, has been very useful in scheduling and blocking the BNR system. This model, which simulates the operation of the systems at a rather macroscopic level works toward optimizing (rather than simply simulating) the rail system. Granting at the outset the major differences between the U.S. system and the British network (e.g., size, one-owner), it would still appear fruitful to consider the usefulness of such a modelling approach in rationalizing rail operations in this country.

Data Systems

This discussion of research needs in rail operation should go no further without alluding to the question of data systems. This is the topic of several other presentations so we need not dwell on it here. However,

as a precondition to improving service, the industry needs to measure what it is doing in rail operations and how well it is performing. Data systems that will allow the industry to perform such measurements are essential. There has been a good deal of progress in this area over the past several years. Among the major efforts are the work in terminal management systems performed under the auspices of the Labor-Management Committee[15] as well as Southern's TPA,[16] Southern Pacific TOPS,[17] and other network control programs. While certainly the industry is in a much better position to measure its service than it was even a short time ago, further refinement of rail data systems is needed. In particular, for our service perspective, the need to measure performance on interline moves is clear. Also of major importance is capturing the portion of the move from shipper's dock to originating yard and from terminating yard to consignee's dock. These local moves are missing from many data systems and yet appear to have a major impact on service reliability.

The above systems reflect the ability of computer and information system technology to measure performance with an eye to isolating problem areas which can then be improved. This is but the first step in the use of computers, however. Research into utilizing the computer to allow the industry to operate in ways inconceivable without the information processing and real time capabilities of the machine is needed. The author believes that the potential of this technology has barely been touched and that service and operating improvements of major impact are within the industry's grasp using this tool.

Cost Models

A final subject of proposed research in the rail operations area is the relationship between rail service quality and operating costs. The

industry needs better ways of understanding the cost implications of various operating alternatives and levels of service. The relationships here are not at all well understood.[18]

One fascinating aspect of this cost question is the notion that improved levels of service (and, in particular, improvements in mean transit time and service reliability) can yield better redistribution of empty cars and deadheading power, less inventorying of empties, better shipper car-detention performance, a shorter car cycle, and therefore, better equipment utilization. Hence, under certain circumstances, one may actually be able to lower costs while providing better service. The relationships described above are shown pictorially in figure 8.

The development of useful cost models is an industry-wide problem. What is emphasized here is that such models be designed in such a manner that the implications of operating changes (and service improvements) can be readily computed.

These cost models, taken together with the service-sensitive demand models described earlier, would give the railroad analyst the ability to evaluate fully the implications of service quality modification from the profitability viewpoint. The relation among the various models is described pictorially in figure 9.

IV-B-2 CAPITAL EXPENDITURES ON RAIL FACILITIES

The potential changes in rail operations described in the previous section are characterized by being implementable within a short period of time. The implicit assumption is that these changes are being implemented within the constraints of existing resources. The area under discussion in this section relates to somewhat longer-term changes. These are changes in the physical facilities available to the rail-operating department,

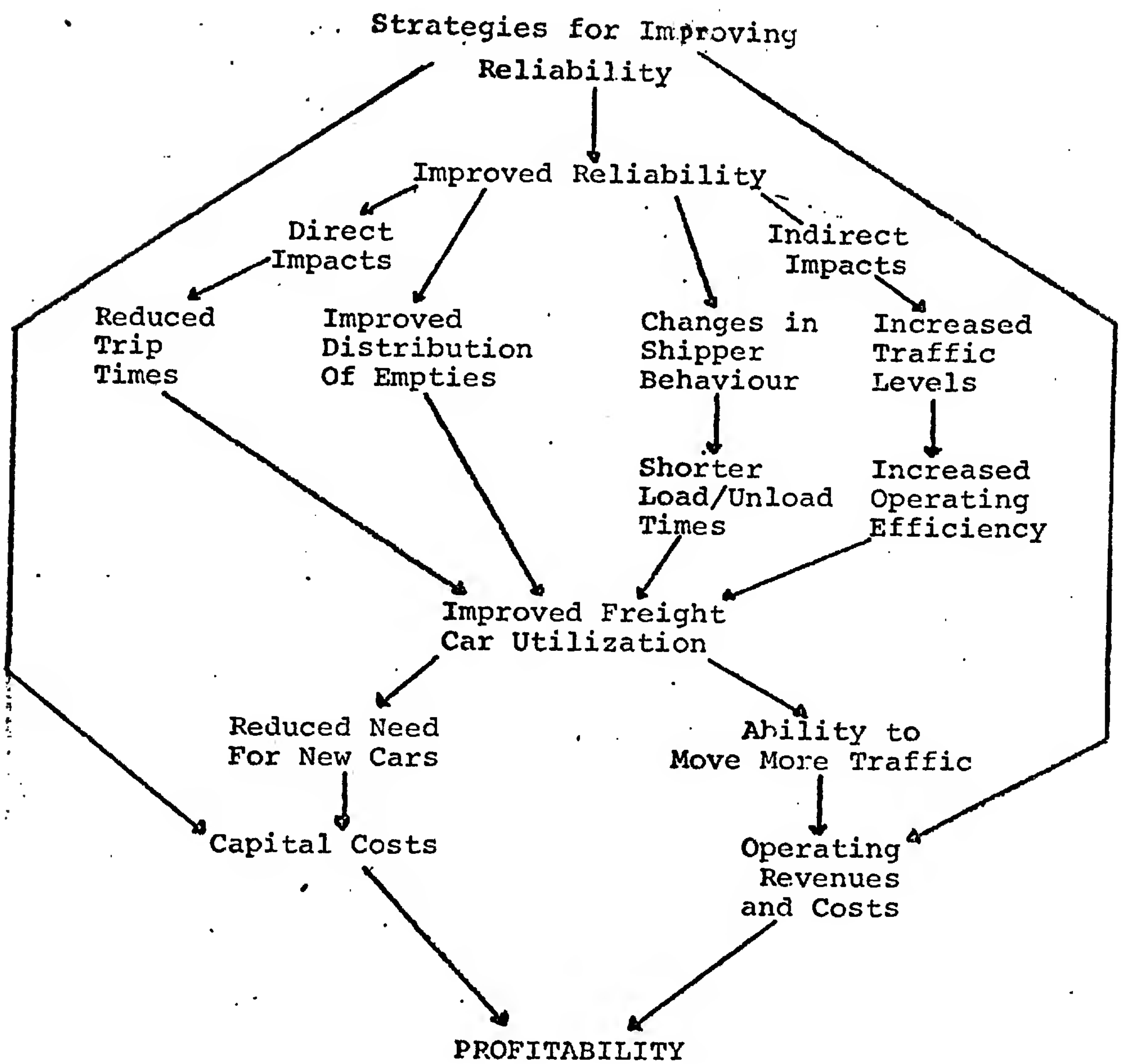


FIGURE 8 A MODELLING FRAMEWORK

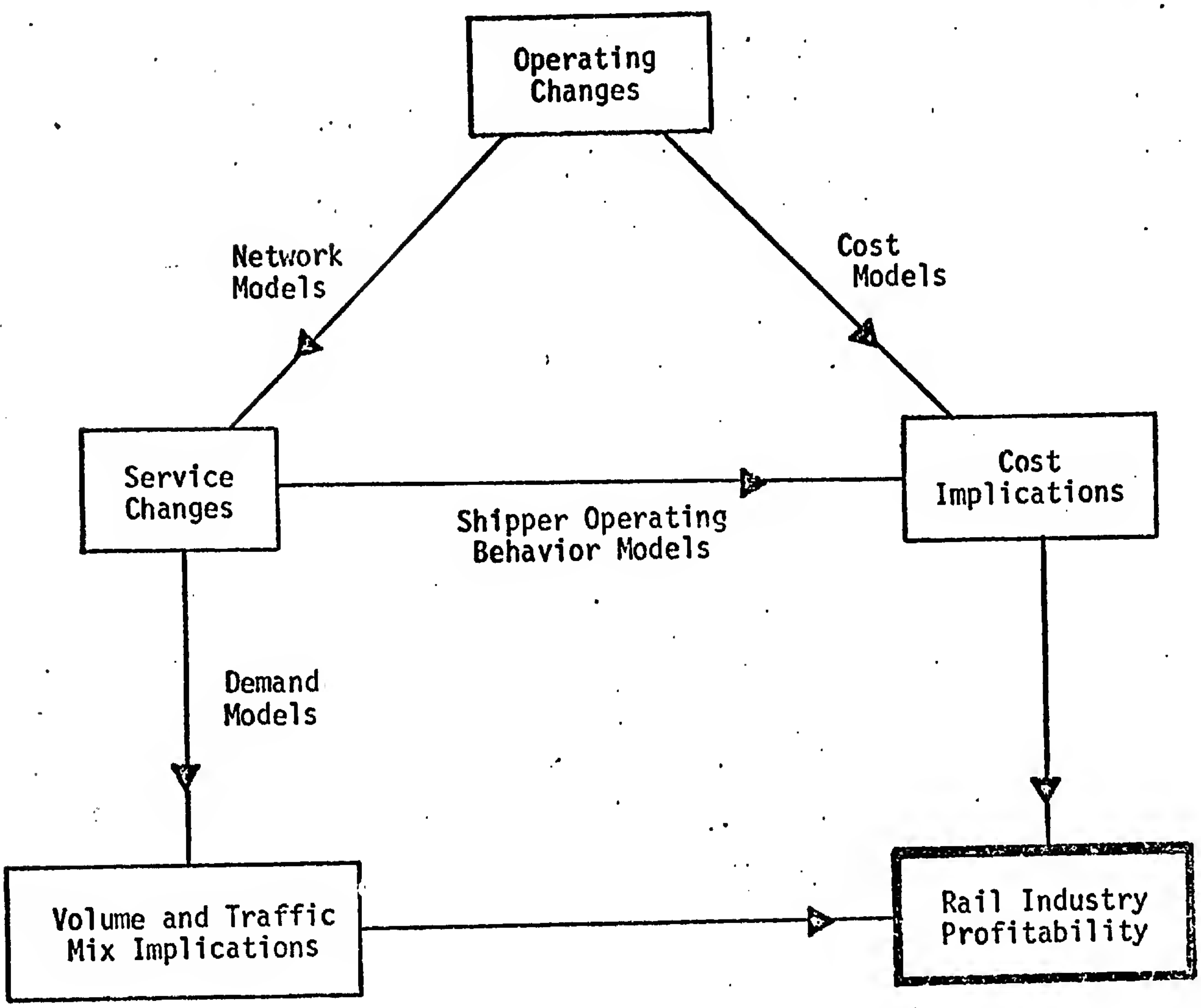


FIGURE 9 IMPLICATIONS OF SERVICE QUALITY MODIFICATION

obtained through a capital investment program. The research area is the development of models for effective investment planning and for providing a useful framework for rail investment planning.

The need is for models that will allow the analyst to predict the impact of particular investment alternatives on service quality and operating costs and hence to be able to make an informed choice among investment alternatives. However, these individual impact prediction models are but a first step. The inherent difficulty in developing such models is compounded by the environment in which investment decisions among alternatives is being made.

1. The investment alternatives are not independent--there are many different investments under consideration at any point in time. Yard construction and upgrading, rolling stock of different types, power, electrification, new track, and structure are but several of the investment possibilities available. A new level of difficulty is introduced when one recognizes that the impact of various alternatives on service and cost are not independent. For example, the impact of the purchase of new power units is related to whether or not a particular terminal is upgraded. How these impacts interact is a difficult modelling issue, as in the complexity introduced by the fact that the combinations of investments can grow to be extremely large.

2. There is a great deal of uncertainty in the system. The railroad investment planner faces the classic case of decision-making under uncertainty. Among the major stochastic elements in the process are:

A. Traffic levels

B. Levels of investment in the future

C. Competitive and regulatory elements

Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9

D. Uncertainty in the prediction abilities of the impact models themselves.

In all, the railroad investment problem is a topic worthy of substantial research. In addition to the development of a variety of impact models (the most notable needs being in the area of yard design and capacity and placement of terminals in the network), a framework for rail investment planning under uncertainty is needed. This framework could usefully build upon the work of Pecknold[19] which deals with multistage investment planning in highway systems.

In summary, investment decisions should be made with service quality in mind. A basic research need is the development of a framework for the understanding of the impacts on service quality of various investment decisions taken in various combinations under conditions of uncertainty with respect to future traffic and investment levels.

IV-B-3 INSTITUTIONAL CHANGES

Operating changes are implementable in short order. Changes derived from capital expenditures take somewhat longer. Still longer term in nature are various institutional changes. These are changes in the basic framework within which the industry operates. By their very nature, institutions were constructed in such a way as to strike a proper balance between various interest groups and any attempt to change (at a particular point in time) "the system" is likely to appear to some group as an attempt to take away some hard-won gains.

Yet, while implementation difficulties are inherent in the process, research into the impacts of proposed changes is appropriate and useful. If the impacts on all actors (industry, labor, shippers, government, etc.) can be understood, trade-offs can be established and compromises reached.

There are any number of institutional issues one could address. For the purpose here, only those with quite direct impact on service and service reliability will be noted.

Work Rules

The relation between operating policy and service quality is clear. Likewise, the relation between operating policies and current operating work rules is strong. A better understanding of how work rules affect operating policies (particularly train length[20]) and how relaxation of particular rules can lead to service improvements is needed.

Network Shape Issues

How the rail network is physically configured and how resources are allocated to this network are closely related to service and service quality. At the same time, network shape is institutionally bound up with the regulatory climate, specifically the I.C.C. Network changes occur very slowly (e.g., through abandonment and mergers), mostly as a result of regulation. A more precise understanding of how service quality relates to network configuration would be useful in illuminating the trade-offs inherent in any network change alternative.

Fractional Per Diem

The rail industry operates on a 24-hour per diem system, with a midnight cut-off. Of current research interest is the possibility of implementing a fractional per diem system. Both 8-hour and hourly systems have been suggested. An understanding of the service implications of these alternatives is required.

Organization Structure in Railroad Companies

The typical railroad company in the U.S. has not changed its organization form for decades. By no means is it clear that the railroad company is optimally structured to address the service needs of its shippers. The organization form of some companies seems more suited to the basically noncompetitive environment of the 1920s and 1930s, an environment that clearly no longer exists.

A fresh look at how a rail freight transportation company should be structured to be able to address today's issues and the issues of the future is needed. Such questions as:

1. How are investment decisions reached?
2. What are the basic information flows in the company and are they the right ones?
3. What is the proper training for a railroad executive?
4. Is the traditional departmental structure of the industry operating, transportation, traffic, etc.) the optimal one?
5. How are level-of-service versus cost trade-offs made? Who makes them?

Much work has been done on organizational structure.[21] Some attempt to utilize these efforts in the rail industry is appropriate.

V SUMMARY

Research needs in the area of service reliability have been described. Two major areas have been discussed. These are the shipper's perspective and the supply of transportation service. Within the first of these is the area of highest priority in the author's view, namely, the development of a better and more quantitative understanding of the relation between service quality and the demand for rail freight service. Also noted here as a lower priority was the issue of the shipper operating responses to rail service and how these are related to equipment utilization and, hence, to rail costs and service.

The supply of transportation service was subdivided into the areas of rail operations, capital expenditures, and institutional changes. setting of priorities here is difficult in that the areas are interrelated and research costs are likely to be very different in size. However, the

author is inclined to set priorities as follows:

Impacts of Network-Wide Operating Changes on Service	(IV-B-1 p. 98)
Impacts of Various Capital Investment Alternatives on Service	(IV-B-2 p. 101)
Data Systems	(IV-B-1 p. 99)
Overall Investment Framework	(IV-B-2 p. 101)
Cost Models	(IV-B-1 p. 100)
Various Institutional Changes	(IV-B-3 p. 105)
Hourly Per Diem	
Work Rules	
Network Shape	
Organization Structure	

A final comment. Service and service reliability is without question an important issue in the rail industry. It is also a researchable question. However, research is not enough. It can only highlight the needs and suggest industry strategies. A continuing and disciplined program of implementation, experimentation, and innovation within the rail industry is required if this research is not to produce a set of dust catchers. By the same token, the research community should discipline itself to produce intellectually sound but pragmatic, implementable, research results.

Based on recent experiences with the interaction of the rail industry and the research community, the author has confidence that a great deal can be accomplished in the future in this area.

BIBLIOGRAPHY

America's Sound Transportation Review Organization. The American Railroad Industry Today: A Prospectus. Washington, D. C., June 1970.

Association of American Railroads. Yearbook of Railroad Facts. Washington, D.C. 1971.

Belovarac, K. and J. S. Kneafsey. Determinants of Line Haul Reliability. Studies in Railroad Operations and Economics, vol. 3. Research Report R72-38, 1972.

Ben-Akiva, Moshe. "Structure of Passenger Travel Demand Models." Ph. D. dissertation, M.I.T., June, 1973.

__________ Marc Terziev and Paul Roberts. "Freight Demand Modelling A Policy Sensitive Approach." Center for Transportation Studies, M.I.T., presented to the 47th National Operations Research Society of America meeting, Chicago, Illinois, April, 1975.

Chandler, A. D., Jr. Strategy and Structure. M.I.T. Press, 1962.

Charles River Associates. The Role of Factor Prices in Transportation Planning for Developing Countries. Study for the Office of International Affairs, U. S. Department of Transportation.

Folk, J. F. "Models for Investigating the Unreliability of Freight Shipments by Rail." Ph.D. dissertation, M.I.T. Operations Research Center, 1972 (unpublished).

__________ Models for Investigating Rail Trip Time Reliability. Studies in Railroad Operations and Research, vol. 5, Research Report R72-40, 1972.

__________ Some Analyses of Railroad Data. Studies in Railroad Operations and Economics, vol. 6. M.I.T. Research Report R72-41, 1972.

__________ A Brief Review of Various Network Models. Studies in Railroad Operations and Economics, vol. 7. M.I.T. Research Report R72-42.

Hartwig, James and William Linton. Disaggregate Mode Choice Models of Intercity Freight Movement. M.S. Thesis, Northwestern University, Department of Civil Engineering, June, 1974.

Jennings, A. A. "The Effect of Train Length on the Reliability of Operation of Freight Yards." S.M. thesis. M.I.T. Department of Civil Engineering, 1972 (unpublished).

Kneafsey, James T. Costing in Railroad Operations: A Proposed Methodology. Studies in Railroad Operations and Economics, vol. 13. M.I.T. Research Report R75-15.

Kolsen, H. M. The Economics and Control of Road Rail Competition. Sydney, Australia: Sydney University Press, 1968.

Kullman, Brian C. "Choice of Mode Between Rail and Truck in the Intercity Freight Market." Ph.D. dissertation, M.I.T. Department of Civil Engineering, 1972 (unpublished).

Lang, A. S. and R. M. Reid. Railroad Car Movement Reliability: A Preliminary Study of Line Haul Operations. Studies in Railroad Operations and Economics, vol. 1. M.I.T. Research Report R70-74.

__________, and C. D. Martland. Reliability in Railroad Operations and Economics. Studies in Railroad Operations and Economics, vol. 8. M.I.T. Research Report R72-74, 1972.

Martland, Carl D. Rail Trip Time Reliability: Evaluation of Performance Measures and Anaylsis of Trip Time Data. Studies in Railroad Operations and Economics, vol. 2. M.I.T. Research Report R72-37, 1972.

__________ Improving Railroad Reliability: A Case Study of the Southern Railway, Studies in Railroad Operations and Economics, vol. 10. M.I.T. Research Report R74-28.

__________ Procedures for Improving Railroad Reliability, Studies in Railroad Operations and Economics, vol. 12. M.I.T. Research Report R74-30.

__________ "Origin to Destination Unreliability in Rail Freight Transportation," Engineer's thesis. M.I.T. Department of Civil Engineering, 1972 (unpublished).

Midwest Research Institute. The AAR Network Simulation System, program documentation for the network simulation model developed for the Association of American Railroads, 1971.

Moody's Investors Corporation. Moody's Transportation Manual. New York: Moody Publishing Co., 1971.

O'Doherty, J. D. "Classification Yard Effects in Rail Freight Movement Reliability." Civil Engineer's thesis, M.I.T., 1972 (unpublished).

Pecknold, Wayne M. "Evolution of Transport Systems: An Analysis of Time Staged Investment Strategies Under Uncertainty." Ph.D. thesis, M.I.T., 1970.

Project Team of Labor/Management Task Force on Terminals. A Program of Experiments Involving Changes in Terminal Operations on the St. Louis Terminal Division Missouri Pacific Railroad 1974 Progress Report.

Railroad Systems and Management Association. The Measure of RAilroad Freight Service. Chicago, Illinois, 1972.

Reebie Associates. "Toward an Effective Demurrage System." Final Report to the Federal Railroad Administration, July, 1972.

Reid, R. M., J. D. O'Doherty, J. M. Sussman, and A. S. Lang. The Impact of Classification Yard Performance on Rail Trip Time Reliability. Studies in Railroad Operations and Economics, vol. 4. M.I.T. Research Report R72-39, 1972.

Roberts, P. O., Jr. "The Logistics Management Process as a Model of Freight Traffic Demand." Harvard Business School Working Paper, HBS71-11, April, 1971

__________, and D. T. Kresge. Techniques of Transport Planning, vol. II. The Brookings Institution, 1972.

ROUTESTRAT. British Railways, 1973.

Sussman, Joseph M. "A Systems Model of the U.S. RAilroad Industry." presented at the ORSA/TIMS meeting, Chicago, April, 1975.

__________, and Carl D. Martland, Reliability in Railroad Operations: Executive Summary. Studies in Railroad Operations and Economics, vol. 9. M.I.T. Research Report R73-4, 1972.

__________. Improving Railroad Reliability: A Case Study of the Southern Railway, Executive Summary. Studies in Railroad Operations and Economics, col. 11. M.I.T. Research Report R-74-29.

Southern Pacific Railway. Total Operations Processing System (TOPS).

Southern Railway. Time Performance Analyzer Manual, 1971.

Transportation Association of America. Transportation Facts and Trends. Washington, D.C., 1971.

NOTES

1. J. Sussman, "A Systems Model of the U. S. Railroad Industry" presented at the ORSA/TIMS National meeting, Chicago, May 1975.
2. Joseph Sussman, Carl D. Martland, and A. Scheffer Lang, Reliability in Railroad Operations: Executive Summary, Studies in Railroad Operations and Economics, vol. 9, M.I.T. Report R73-4, 1972.
3. Carl D. Martland, Rail Trip Time Reliability: Evaluation of Performance Measures and Analysis of Trip Time Data, Studies in Railroad Operations and Economics, vol. 2, M.I.T. Report R72-37, 1972.
4. A. Scheffer Lang and Carl D. Martland, Reliability in Railroad Operations, Studies in Railroad Operations and Economics, vol. 8, M.I.T. Report R72-74, 1972.
5. σ^2_{LH} is the variance of transit time.
6. Robert M. Reid, John D. O'Doherty, Joseph M. Sussman, and A. Scheffer Lang, The Impact of Classification Yard Performance on Rail Trip Time Reliability, Studies in Railroad Operations and Economics, vol. 4, M.I.T. Report No. R72-39, June, 1972.
7. K. Belovarac and J. T. Kneafsey, Determinants of Line Haul Reliability, Studies in Railroad Operations and Economics, vol. 3, M.I.T. Report No. R-72-38, June, 1972.
8. J. F. Folk, Models of Investigation Rail Trip Time Reliability, Studies in Railroad Operations and Economics, vol. 5, M.I.T. Report No. R72-40, June, 1972.
9. Joseph M. Sussman and Carl D. Martland, Improving Railroad Reliability: A Case Study of the Southern Railway, Studies in Railroad Operations and Economics, M.I.T. Report 74-29, 1974.

10. Moshe Ben-Akiva, Marc Terziev and Paul Roberts, "Freight Demand Modelling: A Policy Sensitive Approach," Center for Transportation Studies, M.I.T. presented to the 47th National Operations Research Society of America Meeting, Chicago, Illinois, April, 1975; James Hartwig and William Linton, "Disaggregate Mode Choice Models of Intercity Freight Movement," M.S. thesis, Northwestern University, June, 1974; and Moshe Ben-Akiva, "Structure of Passenger Travel Demand Models," Ph.D. dissertation, M.I.T., June, 1973.

11. Carl D. Martland, Procedures for Improving Railroad Reliability, Studies in Railroad Operations and Economics, vol. 12, M.I.T. Report R74-30.

12. Reebie Associates, Toward an Effective Demurrage System, Report FRA-OE-73-1, National Technical Information Service PB 212-069.

13. Midwest Research Institute. The AAR Network Simulation System, program documentation for the network simulation model developed for the Association of American Railroads, 1971.

14. ROUTESTRAT. British Railways, 1973.

15. Project Team of Labor/Management Task Force on Terminals, A Program of Experiments Involving Changes in Terminal Operations on the St. Louis Terminal Division Missouri Pacific Railroad, 1974 Progress Report.

16. Time Performance Analyzer Manual, Southern Railway, 1971.

17. Total Operations Processing System (TOPS), Southern Pacific Transportation Co.

18. J. T. Kneafsey, Costing in Railroad Operations: A Proposed Methodology, Studies in Railroad Operations and Economics, vol. 13, Research Report R75-15, 1975.

19. Wayne M. Pecknold, "Evolution of Transport Systems: An Analysis of Time Staged Investment Strategies Under Uncertainty," Ph.D. dissertation, M.I.T., 1970.

20. A. Scheffer Lang and Carl D. Martland, _op. cit._

21. Alfred D. Chandler, Jr., Strategy and Structure, M.I.T. Press, 1962.

CURRENT APPROACHES TO TRAVEL DEMAND FORECASTING*

Marvin L. Manheim**

1. Introduction

One purpose of this paper is to present a brief overview of the major methods now in use in the U.S. for forecasting future passenger travel in urban areas. Many of these methods have been applied also to forecasting intercity and national passenger travel and to predicting freight movements as well.

In the next section, we summarize some of the issues currently being addressed in urban transportation planning studies in the U.S. Then we review briefly the basic theory of transportation systems analysis underlying travel forecasting methods. In the fourth section, we summarize the major alternative approaches available today and show their interrelations in section 5. Examples of two major approaches are given in section 6. The last section presents our conclusions.

2. The Role of Travel Forecasting Models in the Analysis of Transportation Plans and Policies

Now, more than ever before, a multiplicity of urban and regional objectives is leading to the proposal of a wide range of policies to control and shape travel patterns. The number, range, and complexity of these proposed policies is unprecedented. In order to appraise the

---

*This paper draws upon numerous articles and reports to which many other individuals have contributed. Where specific acknowledgement is practical, appropriate references have been cited. The author particularly wishes to acknowledge the contributions of several individuals whose individual and collaborative work, and words, are infused throughout this paper, Moshe Ben-Akiva, William Jessiman, Steven Lerman, Wayne Pecknold, and Earl Ruiter, who could almost be listed as co-authors.

**Professor, M.I.T.

advantages and disadvantages of these proposals, it is essential to be able to predict how travel patterns will be affected.

Among the issues facing urban transportation planners are these:

1. air quality
2. energy
3. land use policy
4. urban congestion
5. transportation impacts on special groups, such as: elderly, handicapped, low-income, etc.
6. public transportation financing and new technology development

In response to these issues, a wide variety of transportation options are being considered. Some of these provide new auto or transit services, while others are disincentives as, for example, steps to reduce auto use:

1. new highway facilities
2. new transit service offerings such as rapid transit extensions, feeder transit service coordination, express bus, subscription bus, jitney, dial-a-ride, taxi, dual-mode, light rail transit, etc.
3. taxes on gasoline
4. gasoline rationing
5. gasoline allocation schemes
6. parking restrictions, bans, or surcharges

7. vehicle exclusion policies--areas, days of the week, time of the day, registration sticker schemes
8. time bans on driving on particular days
9. car pooling programs and incentives
10. restrictions on auto ownership
11. taxes on vehicle ownership
12. congestion tolls
13. staggered work hours
14. parking changes
15. changes in transit fares and subsidies
16. ramp metering and other improvements to highway facilities
17. vehicle emission controls
18. land use controls

In developing a regional or subregional transportation plan for a particular urban area, policy components such as these are combined into overall programs.

The challenge of travel forecasting methods is to be able to predict the responses of potential travellers to such elements. For example, consider these questions on the impacts of various policies designed to reduce air pollution from transportation: How will trip-makers respond to a forced reduction in employee parking at factories and offices? How will they respond to an early morning on-street parking ban, such as the one scheduled for Boston in the spring of next year? Will the result be a decrease in

trip-making? A change of destination? A shift to public transportation, to car pools, or to late arrival (i.e., staggered work hours for employees)? What would the imposition of selective parking surcharges on existing fees do to demand by the various modes? What is the impact of the imposition of tolls on certain facilities? Would prohibitions on entry by private autos into certain areas have a similar effect? Do these policies have undesirable differential impacts on different segments of society? What are the short-term and long-term shifts in residence, location, choice of work place, or household automobile ownership level that will occur as a result of various such air pollution control strategies?

Questions such as these must be answered before a rational policy for reducing air pollution can be selected. Understanding travel demand behavior plays a critical role in arriving at these answers.

3. The Theoretical Basis of Travel Forecasting

Modern travel forecasting methods are based on the theory of transportation systems analysis. This theory has emerged from several sources.[1] In outline, the problem of predicting the flows in a transportation system is a simple application of economic theory: the flows which will result from a particular transportation system (T) and pattern of socioeconomic activities (A) can be determined by finding the resulting equilibrium in the transportation market. If:

V = volume of flow

L = level of service experienced by that volume

F = (V,L) = flow pattern

Then, we find equilibrium by establishing a supply function (S) and demand function (D), and solving for the equilibrium flows (F_o) consistent with both relations:[2]

(1.1) $L = S(V,T)$

$V = D(L,A)$

$F_o = (V_o, L_o)$

(See figure 1, Basic Theory)

While simple in outline, the application of the theory becomes complex in practice for several reasons:

1. The consumer considers many service attributes of the transport system when making a choice (e.g., line-haul travel time, transfer time, walk distance, out-of-pocket cost, privacy, etc.), and thus, L must be a vector with many components
2. Determining the demand functions (as well as other elements) to use is difficult
3. The equilibrium occurs in a network, where flows from many origins to many different destinations interact, competing for the capacity of the network; and the form of these interactions is affected by the topology of the network

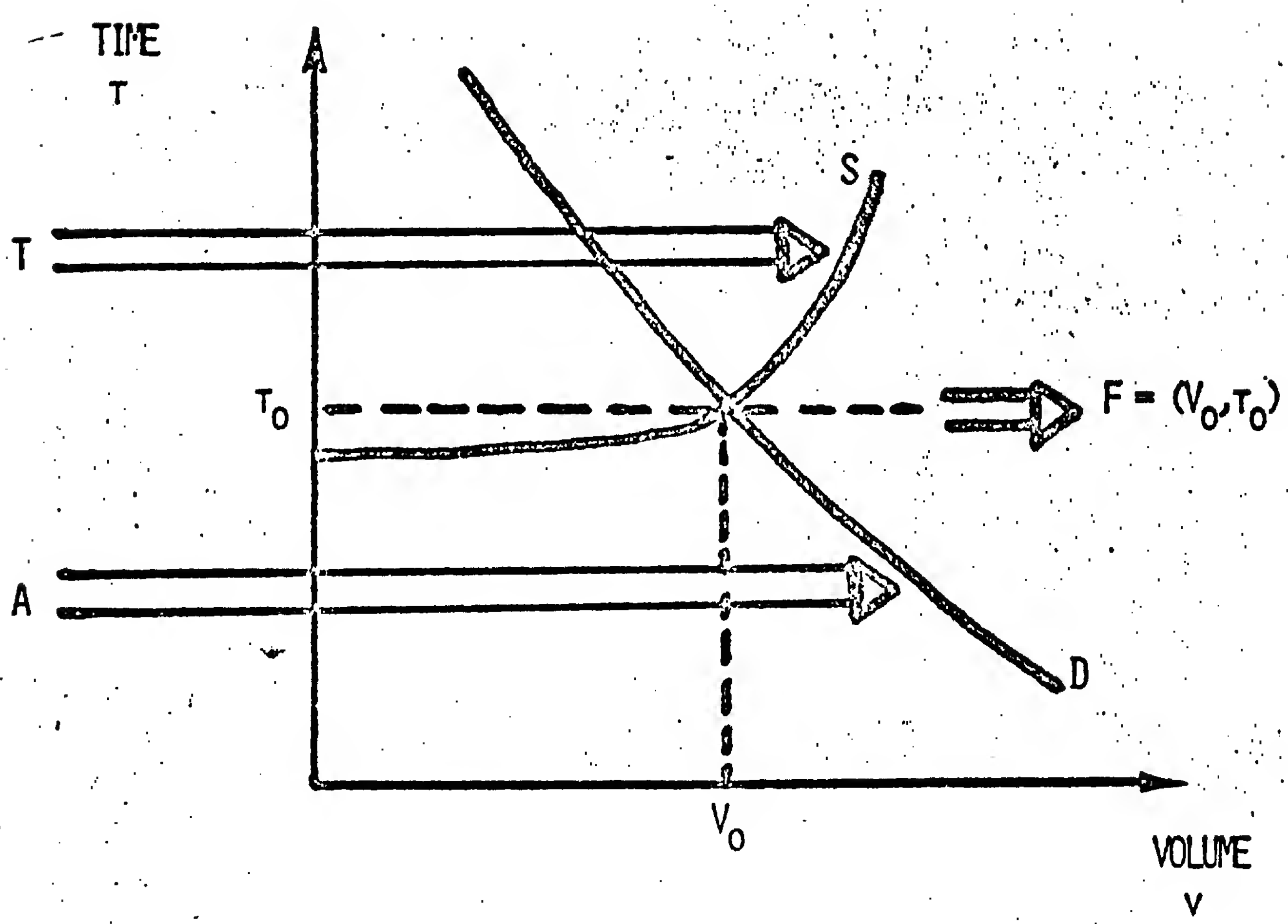


FIGURE 1. BASIC THEORY

Thus, fairly elaborate computational schemes are required to actually determine the equilibrium flows F_o for a particular (T,A).

In the case of a multimodal network, the symbol V represents an array of volumes:

(1.2) $V = \{V_{kdmr}\}$, for every k,d,m and r,

where V_{kdmr} is the volume flowing from origin zone k to destination zone d via mode m and path r of that mode, and the brackets { } indicate a set of elements V_{kdmr}. Ideally, once we have established our demand and supply functions, we would then like to be able to turn directly to an equilibrium-calculating procedure to "solve" the two sets of relationships to find the equilibrium flow pattern. The result of this computation would be the two arrays comprising that flow pattern:

(1.3) $F_o = (V_o, L_o)$, where

$V_o = \{V_{kdmr}\}$ for every k,d,m, and r; and

$L_o = \{L_{kdmr}\}$ for every k,d,m, and r.

In words: we should get out of our equilibrium procedure the volumes, and the levels of service experienced by those volumes, from k to d by mode m and path r.

Unfortunately, at this state of the science of transportation modelling, while several systems of transportation models exist, there is not even one operational model which solves for these equilibrium flows exactly and directly.[3] There are available a number of different systems of models, each of which represents a

different operational approach to computing equilibrium in transport networks. The differences in these approaches are reflected both in the computational algorithms and in the structure of the demand models which are used.

One particular computational scheme is that which has been used historically in urban transportation planning studies. In this approach the equilibrium flows are estimated in a sequence of steps, commonly called trip generation, modal split, and traffic assignment.[4] Correspondingly, the demand function, D, is represented as a sequence of functions: trip generation (and attraction) equations, trip distribution procedures, modal split equations, and the minimum-path rules of the traffic assignment procedures. We will refer to this approach as the UTMS--the Urban Transportation Model System.

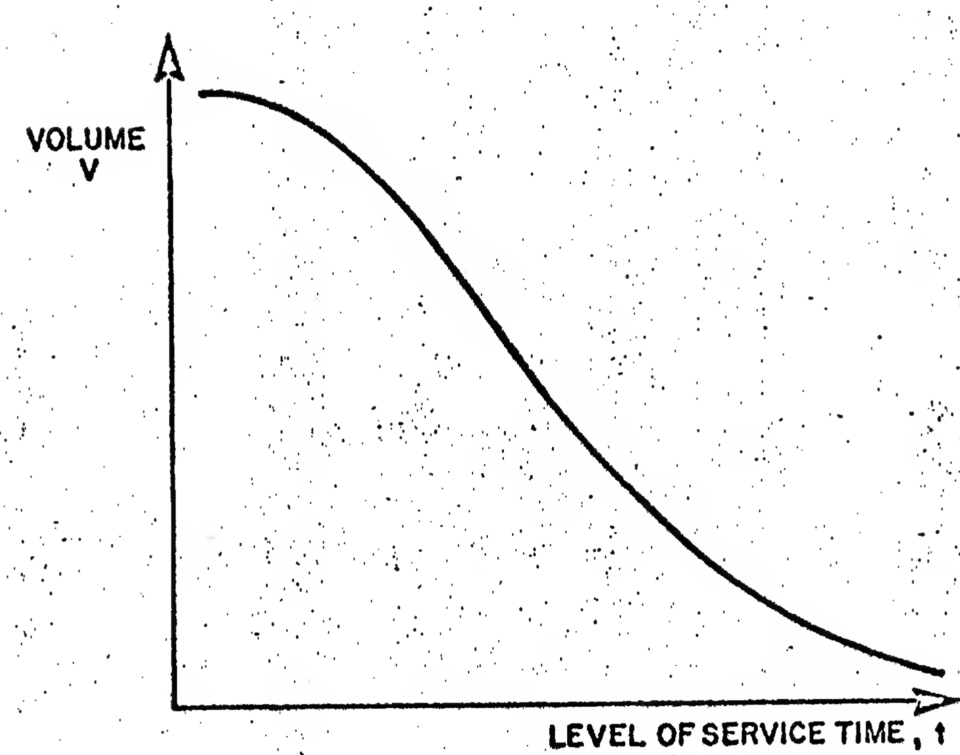
More recently, other alternative approaches have been developed. In the following sections of this paper, we will summarize the major alternative approaches which now exist.

4. The Major Alternative Approaches

In the preceding sections, the basic concept of a demand function was introduced as a way of representing consumer behavior. A wide variety of different types of demand models can be developed. The objective of this section is to present some of the most important distinctions among the types of demand models used in travel forecasting in order to illustrate the variety of modelling choices available.

FIGURE

The Demand Function



In this discussion, we will tend to emphasize primarily "structural" or "abstract" features of the various choices. However, we must never forget that the basis of everything we do in demand modelling is based on trying to represent the behavior of the consumer. To emphasize this, we will relate the discussion of modelling choices to the way in which the choices open to the individual consumer are assumed to be perceived and operated upon.

In general, in urban transportation, there is a wide range of choices open to a traveller. These choices are among combinations (f,d,m,r,h,ao, res, emp), where:

f = frequency of trips

d destination

m = mode

r = route

h = time of trip

ao = number of autos owned

res = residential location

emp = employment location

For most of urban travel forecasting to date, it has been assumed that (res) and (emp) are fixed for each traveller (or alternatively, an urban growth and land use model is used to predict these long-run effects separately from the short-run travel decisions), and so travel forecasts deal primarily with choices of (f,d,m,r,h,ao).

Further, in the past it has also been usual to separate choices of time of trip and of auto ownership from the other choices, reducing the primary focus of travel demand models to (f,d,m,r).

Looking at the behavior of groups of consumers, we are thus concerned with predicting the aggregate total volume choosing a particular combination, in this form:

V_{kdmr} - the volume of trips from origin zone k to destination d by mode m and path r

Note that we have specified the origin of the trips (k), and that frequency has been replaced by the total number of trips made (i.e., frequencies of one, two, or more trips per individual per time period).

There are a wide variety of different ways in which specific demand models might be constructed for use in transportation analyses. In this section, we will discuss some of the major differences in types of demand models: variables included, functional forms, and three structural features which will be our primary concern: aggregate versus disaggregate, probabilistic versus deterministic, and simulteneous versus sequestial.

4.1 Variables Included

Recall that there are two basic sets of variables, activity system variables (A) and level-of-service variables (L) in the demand function (1.1). One important set of choices concern what activity system and level-of-service variables to use. The activity system may be described in terms of such variables as population, employment, income, household size, stage in family-life-cycle, etc. The level of service of the transportation system could be described in terms of travel time, separated into in-vehicle and excess time, time reliability, service schedule, out-of-pocket cost, perceived security, tolls, etc.

4.2 Functional Forms of the Models

A wide variety of alternative functional forms are possible. The most common are the product, linear, exponential, and logistic or logit forms.

4.3 Aggregate and Disaggregate

Recall that we are typically dealing with the consumer's choice of the set of travel options (f,d,m,r). If we consider an individual traveller (or household) i at location k, we can ask, which combination (f,d,m,r) does this individual pick?

In this case, our demand function is disaggregate; it predicts the choice of a single decision-making unit (either individual or household). We can write the demand function as

(4.3-1)
$$X_{fdmr} = g_i\ (A,L)$$

where: $$S_{fdmr} = \begin{cases} 1 \text{ if } i \text{ chooses } (f,d,m,r); \\ 0 \text{ if } i \text{ chooses some other combination} \end{cases}$$

Let us now consider a group of individuals, e, which we will define as a "market segment." (Generally, we try to group individuals or households into market segments such that the travel behaviors of all the individuals in a single market segment are relatively similar.)

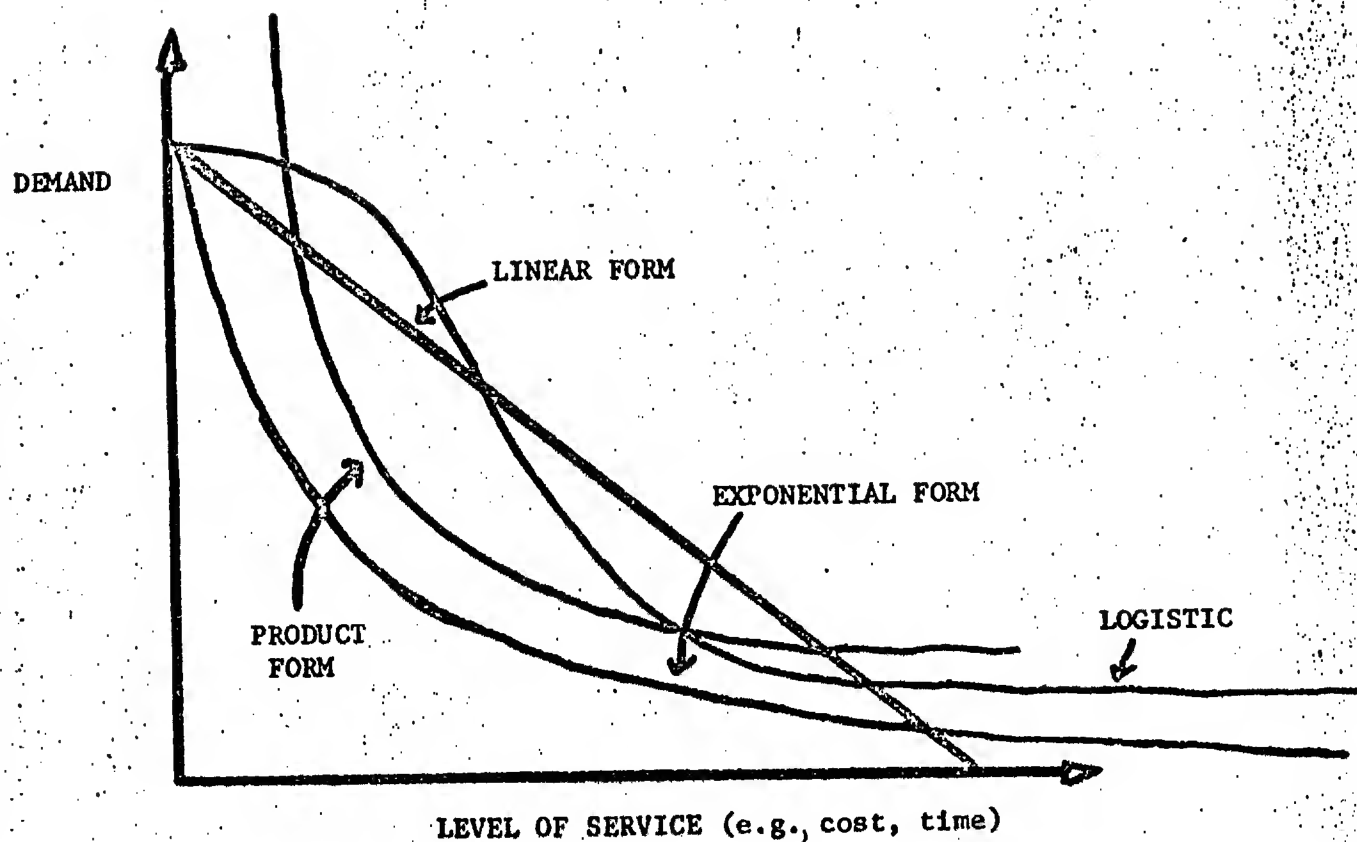
In this case, we write:

(4.3-2) $$V^e_{kdmr} = f_{ek}\ (A,L)$$

This aggregate demand function gives the number of individuals in market segment e at location k who will choose (d,m,r).

FIGURE 4.2-1

ALTERNATIVE FORMS OF DEMAND FUNCTIONS



4.4 Probabilistic and Deterministic

A probabilistic demand model gives an estimate of the relative likelihood of different choices being made. A deterministic demand model, on the other hand, predicts that one and only one specific choice will be made.

Usually, disaggregate travel demand models are also formulated as probabilistic models. If we consider an individual traveller or household i, we can ask, what is the probability p_i that this individual at location k will choose a particular combination of (f,d,m,r)? In general, we know this will be a function of some of the social, economic, and other characteristics of that individual (A_i), the characteristics of the activities which could be undertaken at various destinations d, (A_d), and the level-of-service characteristics L_{kdmr} for each path r in a mode m from location k to destination d:

$$(4.4\text{-}1) \quad P_i(f,d,m,r) = g_i[\{A_i, A_d, L_{kdmr}\} \text{ for all } d,m,r]$$

We write this as a probability in part because there will likely always be some inherent randomness in each individual's decision process.

Usually, aggregate models have been assumed to be deterministic, although a probabilistic form would also be reasonable:

$$(4.4\text{-}2) \quad P_{ke}(V^e_{kdmr}) = f_e\ (A,L)$$

4.5 Simultaneous and Sequential

In analyzing the travel behavior of any individual or group or individuals, a number of alternative assumptions about the

pattern of behavior can be made. One assumption of particular importance is whether individuals make simultaneous or sequential choices.

At the disaggregate level, we can write the two cases as follows:[5]

(4.5-1) Sequential: $p_i(f) \cdot p_i(d|f) \cdot p_i(m|f,d) \cdot p_i(r|f,d,m)$

(4.5-2) Simultaneous: $p_i(f,d,m,r)$

These two equations are mathematically equivalent only if, in (4.5-1) $p_i(f)$, the probability of a given level of trip frequency (f) is indeed independent of destination (d), mode (m), and route (r). Similar conditions must hold for $p_i(d|f)$, etc. If these conditions are met, then the following behavioral interpretation can be given to (4.5-1): the potential traveller chooses first how many trips to make, then a destination, then (conditional on choice of destination and frequency) a mode, and then (conditional on having made all the other choices), a route. If these conditions (or analogous conditions for an alternative sequence of decisions), are not met, then (4.5.2) can be chosen. The corresponding behavioral interpretation is that the individual chooses his trip frequency, destination, mode, and route simultaneously.

Alternative estimation (model calibration) approaches are available for each case. Estimation of a simultaneous model, such as (4.5-2), is appropriate regardless of any behavioral choice sequence which may exist. Once such a simultaneous model has been estimated, any desired marginal and conditional probability functions can be derived using the basic formulas of probability theory--for example, a model split probability function, conditional on choice

of destination and trip frequency as in (4.5-1): $p(m|f,d)$. However, since simultaneous models are harder to estimate, information on the actual sequence of choices, if available either from prior empirical studies or from behavioral theories, can be used to simplify estimation by using sequential models. If the proper sequence is assumed, the results should be identical, since the two cases are mathematically equivalent and, thus, either can be converted to the other, using the mathematics of probability theory. If the sequence assumed as a basis for estimation is not correct, this will become evident in one or both of the following ways:

1. If both types of models are estimated, they will not obey the relationships required by probability theory. Ben-Akiva (1973) has demonstrated this case.
2. If only a simultaneous model is estimated, it will not be possible to derive sequential submodels with the necessary independence properties.

The most general assumption is that the decision process of the individual traveller is a simultaneous one and, therefore, we should estimate a simultaneous model. Then, any desired sequence of models can be derived.

Although from a behavioral point of view it would seem best to begin with an assumption of a simultaneous choice model (where there are no behavioral arguments for any particular sequential model), there may be practical problems. Because of the larger number of possible combinations of choices, the number of explanatory variables and allowable interactions between variables may provide some difficulty,

and the model may become very complex. However, these problems could be resolved in any of several ways other than by assuming a sequential model, such as reducing the number of attributes a traveller is assumed to consider or the number of potential destinations.

A similar distinction can be made at the aggregate level:

(4.5-3) Sequential:

$$V_k = f_1(A_k, D_d, L_{kdmr})$$

$$V_{kd} = f_2(A_k, D_d, L_{kdmr}, V_k)$$

$$V_{kdm} = f_3(A_k, D_d, L_{kdmr}, V_{kd})$$

$$V_{kdmr} = f_4(A_k, D_d, L_{kdmr}, V_{kdm})$$

(4.5-4) Simultaneous:

$$V_{kdmr} = f(A_k, D_d, L_{kdmr})$$

Again, we note that the sequence in (4.5-3) is one of many possible sequences.

It is important to note that the sequence in (4.5-3) is the same general form as in the traditional four-step demand models used in urban transportation and introduced in section 3--The Urban Transportation Model System (UTMS). Recent theoretical results, including the development of a "General Share Model" [12], show that any desired sequence, such as (4.5-3), can be derived from a given simultaneous form (4.5-4); and that, under certain conditions,[6] an equivalent simultaneous form (4.5-5) can be derived from the sequential form (4.5-3). (The logic of these results is very similar to the basic mathematics of probability theory which apply at the disaggregate levels.) Again, however, in each case it is likely that equivalent forms (the two simultaneous forms or the

two sequential forms) obtained in the two different ways (direct estimation or mathematical derivation of one from the other) will be different.

Therefore, the most general approach at the aggregate level is also to assume a simultaneous choice form, estimate that form, and then derive analytically the corresponding desired sequential forms, just as was suggested for the disaggregate models above.

In this way, some of the objections to the conventional four-step models (cf. section 6.1.5 below) can be overcome; since level-of-service variables will appear in each step, there can be explicit behavioral structure, and valid estimation procedures can be used.

4.6 Other Modelling Choices

In addition to those explicitly described in this section, other key modelling choices include: choice of market segments; commodity-dependent versus commodity-independent attributes (e.g., "abstract mode"); incorporation of activity variables explicitly or via stratification. These are, in general, choices which must be made by examination of the results of statistical testing of hypotheses reasoning.

5. Relationships of Major Model Development Alternatives

The preceding discussion suggests a bewildering variety of combinations of modelling choices. In this discussion, we will focus on the three sets of structural relationships: aggregate versus disaggregate; probabilistic versus deterministic; simultaneous versus sequential. There are at present four basic alternatives for model estimation in actual use:

1. Aggregate, deterministic, simultaneous models--Charles River Associates San Francisco direct demand model (CRA,1967), for example
2. Aggregate, deterministic, sequential models--the traditional UTMS models, introduced in section 3, for example
3. Disaggregate, probabilistic, simultaneous models--as estimated by Ben-Akiva (1973), for example
4. Disaggregate, probabilistic, sequential models--the bulk of the behavioral disaggregate modal split models [Stopher, 1969; Stopher and Lisco; Quarmby; Reichman and Stopher; Charles River Associates (1972); McFadden]

In addition to these four alternative forms, there are also four basic ways to use these estimated models in forecasting. Because we must forecast total trips made (by various modes to various destinations, etc.) rather than probabilities of trips made, disaggregate models must be aggregated in the forecasting process.

Prediction of equilibrium flows requires an equilibration method used with an aggregate demand model (either estimated in aggregate form or constructed from an estimated disaggregate model by aggregation). If the resulting aggregate model is simultaneous, then equilibrium can be computed in a single step with a "direct" approach. If the demand model is sequential, i.e., a sequence of equations, then equilibrium is usually computed in a sequence of steps, in an "indirect" approach. These two choices for forecasting are indicated as E and F in figure 6-1. The four basic alternative ways to estimate <u>and</u> use models for forecasting as shown in figure 6.1 are:

1. Estimation of an aggregate, deterministic, simultaneous model for all aspects of trip choice; and using that same model for forecasting in a direct approach. (This is represented by line CE in the figure.)
2. Estimation of a disaggregate, probabilistic, simultaneous model of all aspects of trip choice, and then using an aggregation process to obtain an aggregate simulteneous model for forecasting in a direct approach. (This is represented by line AE in the figure).
3. Estimation of an aggregate, deterministic, sequential model of trip choice, by estimating each component in the sequence, and then using each of the aggregate model components separately for forecasting in an indirect approach. (Represented by line DF in the figure.)
4. Estimation of a disaggregate, probabilistic, sequential model, also estimating each component separately, then using an aggregation process to separately aggregate each component up to a sequential aggregate model, and then using the aggregate model components for forecasting, in an indirect approach. (Represented by line BF in the figure.)

In the past, these four approaches have seemed to be essentially different and competitive. However, recent research results show that these approaches are all interrelated; they are different views of the same travel behavior. Once this is understood, relations among the different views can be developed and, thus, additional approaches become available.

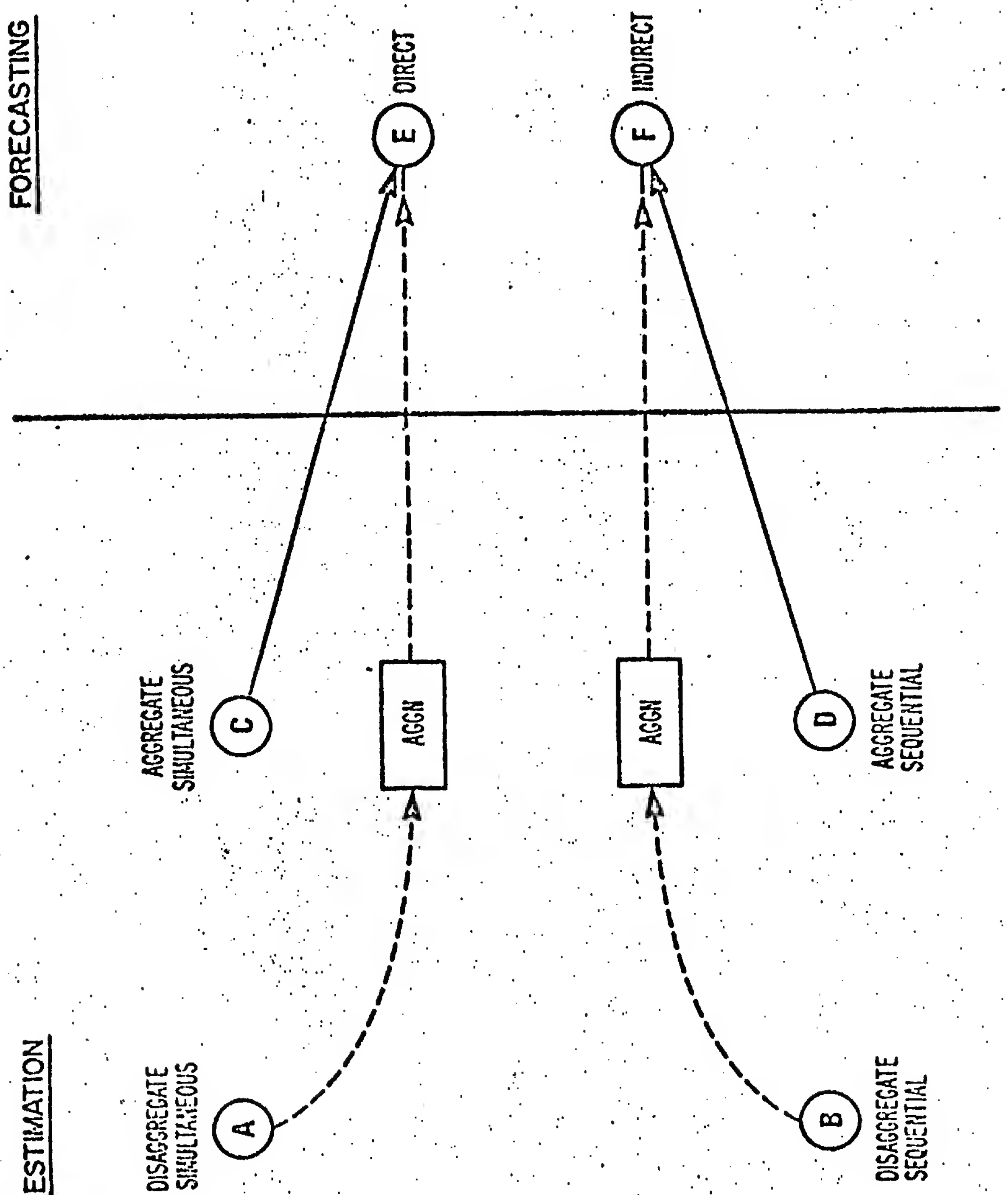


FIGURE
FOUR BASIC ALTERNATIVE DEMAND MODELLING APPROACHES

The first step in laying out these interrelations is to note the connection between simultaneous and sequential forms at the disaggregate level, that of the individual consumer. (The basic theory was developed by Ben-Akiva (1973). It is possible to go from simultaneous disaggregate to sequential disaggregate models, using the basic mathematics of probability theory; and under specific conditions to go from sequential to simultaneous disaggregate models.

The second step is to note the connection between simultaneous and sequential forms at the aggregate level, that of a "market segment" composed of a group of individual consumers. The basic theory here was developed in the concept of the General Share Model (Manheim, 1973). This development shows how it is possible to go from simultaneous aggregate ("explicit") to sequential aggregate models, using relations very analogous to those applying at the disaggregate level; and under special conditions to go from sequential aggregate to simultaneous aggregate models.

The third step is to note the connection between aggregate and disaggregate forms. A general theory of aggregation is not yet available; however, assumptions can be made which lead to directly useable practical procedures (Koppelman, 1975).

Once these relationships are recognized, additional demand modelling approaches become available, as shown in figure 6-2. For example, an aggregate simultaneous model can be estimated directly or by aggregating a disaggregate simultaneous model. From this can be derived an aggregate sequential form of use in an indirect forecasting approach, such as the four steps used in the traditional UTMS

and embodied in numerous software packages. This is represented by lines CF and GF respectively in Figure 6-2. Similarly, for the sequential model, we can derive a simultaneous form of that model for forecasting in a direct approach. In Figure 6-2, this is represented by lines BHE and DE respectively. So, for example, in recent work for CALTRANS (Nestle and Young) the present California generation, distribution, and mode split relations were combined into a single simultaneous model.

Thus, we now have a great number of ways to estimate and use demand models (eight ways, structurally, as shown in Figure 6-2).

The benefit of this wide range of alternative combinations of estimation and forecasting techniques is that we can exploit the data efficiency of disaggregate models, and produce models similar to those conventionally used today, but which are behaviorally more valid for use in policy-oriented studies.

Moreover, because of these relationships which exist between alternative estimation methods and alternative forecasting methods, we can base the choice of forecasting methods on grounds solely of computational efficiency. We are thus free to choose behavioral assumptions and estimation methods without concern for which form will be used for forecasting:

1. the "software" for forecasting can be designed to operate either with the conventional 4-step models or with simultaneous models;
2. disaggregate methods can be used for efficiency in model development;
3. aggregation of disaggregate models can be done to produce both simultaneous and sequential forms, to allow the user

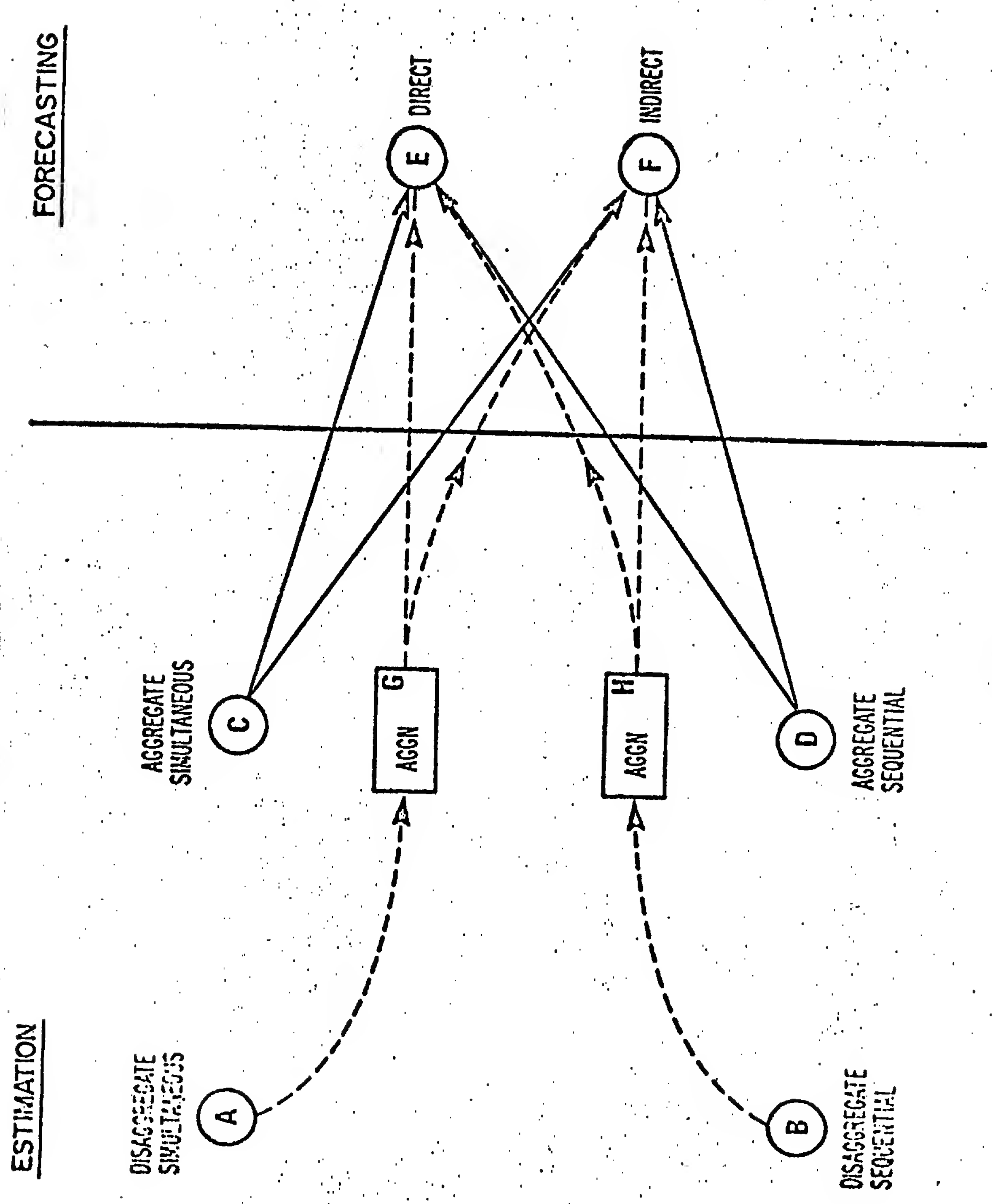


FIGURE
EIGHT TOTAL ALTERNATIVE DEMAND MODELLING APPROACHES

of the models to choose which ones are most useful in a specific forecasting task.

6. Examples of Major Demand Modelling Choices[7]

In the preceding section, the set of combinations of major demand modelling choices which have been chosen so far for demand model development was identified below:

Group I - aggregate, sequential, and deterministic

Group II - aggregate, simultaneous, and deterministic

Group III - disaggregate, sequential, and probabilistic

Group IV - disaggregate, simultaneous, and probabilistic

In the following sections, we will give examples of two of these groups: I and IV.

6.1 The Traditional Four-Step Approach: Aggregate Sequential Models

To review: The basic choice open to prospective travellers is among combinations (f,d,m,r,h;ao,res,emp)

where: f = frequency of trips[8]

d = destination

m = mode

r = route

h = time of trip

ao = number of autos owned

res = residential location

emp = employment location

For most urban travel forecasting, it has been conventional to assume that residential location (res) and employment (emp) are fixed for each traveller (or alternatively, an urban growth and land use model

has been used to predict these "long-run" effects separately from the "short-run" travel decisions). Thus, models have dealt primarily with choices of (f,d,m,r,h;ao). Furthermore, typically the choices of time of travel and of automobile ownership are usually dealt with separately, leaving (f,d,m,r) as the primary concern of travel models.

Usually, we are interested in the behavior of groups of individuals. A market segment is defined as a group of individuals (or households) with similar travel behavior. Then, if we consider all of the individuals in market segment e who reside in zone k of an urbanized area, the demand function can be described as:

$$(6.1\text{-}1) \qquad V^{e}_{kdmrh;ao} = f(A,L)$$

where A,L are a variety of relevant activity system and servica variables describing the choices available; and $V^{e}_{kdmrh;ao}$ is the total number of trips made by members of market segment e in origin zone k owning "ao" automobiles, to destination d by mode m and route r at time h.

In the typical approach, where it is assumed that choices of time of travel and of auto ownership are made separately from the others, this becomes:

$$(6.1\text{-}2) \qquad V^{e}_{kdmr} = f(A,L)$$

Historically, this demand model has been broken into four components:

$$(6.1\text{-}3) \qquad V^{e}_{kdmr} = f_1(A,L) \cdot f_2(A,L) \cdot f_3(A,L) \cdot f_4(A,L)$$

where f_1 = trip generation submodel

f_2 = distribution submodel

f_3 = modal split submodel

f_4 = route choice ("network assignment") submodel

By breaking the demand function into four submodels, it was possible to break the forecasting of future travel into a sequence of four steps:

trip generation:

V_k^e = total volume of trips generated by zone k in market segment e

(6.1-4) $V_k^e = f_1(A,L)$

trip distribution:

V_{kd}^e = total volume of trips originating in zone k and travelling to destination zone d, in market segment e

(6.1-5) $V_{kd}^e = f_2\ (A,L) \cdot V_k^e$

model split:

V_{kdm}^e = total volume of trips going by mode m from origin zone k to destination zone d in market segment e

(6.1-6) $V_{kdm}^e = f_3\ (A,L) \cdot V_{kd}^e$

network assignment:

V_{kdmr}^e = total volume of trips going by route r of mode m from origin zone k to destination zone d in market segment e

(6.1-7) $V_{kdmv}^e = f_4\ (A,L) \cdot V_{kdm}^e$

In equations (6.1-5, 6.1-7), the functions f_2, f_3, and f_4 represent "share" functions--they serve to split a larger trip total into a number of components: In equation 6.1-6, for example, f_3 is applied to the total trips from k to d by segment e to obtain the share of these trips which will go by each of the available modes, m.

It is instructive to examine each of these submodels in detail.

6.1.1 Trip Generation

Trip generation is the first sequential step, involving the prediction of total trips from an origin or to a destination by trip purpose (7). The functional form is usually linear. Symbolically,

$$V_k^e = a^e + {}_i b_i A_{ki} + {}_i C_i^e f_i(L_{k\cdot})$$

where: V_k^e = trips of purpose e generated in origin k

a^e, b_i^e, c_i^e = empirical parameters

A_{ki} = activity system variable i for origin k_i

$f_i(L_{k\cdot})$ = function of level of service variables from k to all destinations

Typically, activity system variables used are average annual income, average number of autos owned, number of workers per household, percentage of households having an income greater than a specified value; zonal population, acres of land in various land-use categories and zonal employment. A typical level-of-service function is accessibility, defined as a function of travel times or distance; but such functions are seldom used in trip generation.

Example

The two trip-generation equations which follow are typical of equations used in the traditional four-step approach:

$$\frac{\text{Total Home-Based Trips}}{\text{Dwelling Unit}} = 0.69 + 1.94\ (\text{cars/D.U.}) + 1.39\ (\text{residents} \geq 5\ \text{yrs./D.U.})$$

$$\begin{aligned}\text{Home-Based Work Trips} = {} & 1.148\ (\text{workers}) \\ & + 0.569\ (\text{households}) \\ & + 0.019\ (\text{residents} \geq 5\ \text{years}) \\ & - 0.144\ (\text{residents} \times \text{distance from CBD}) \\ & + 0.488\ (\text{residents})\end{aligned}$$

Note that neither equation includes any level-of-service variables: zonal trips will therefore be predicted to remain constant, no matter what level of transportation service is provided.

6.1.2 Trip Distribution

The second sequential step in the conventional four-step approach is trip distribution, the prediction of trips from origin to destination. The independent variables are the "trip ends" resulting from the previous step, plus level-of-service variables. Symbolically,

$$V^e_{kd} = f_n(V^e_k, V^e_d, L_{kd}) \qquad (6.1\text{-}8)$$

where V^e_{kd} = trips of purpose e from origin k to destination d

V^e_k, V^e_d = results of the trip-generation step

L_{kd} = level-of-service variables between k and d

The two most common functional forms are the gravity model and the opportunity model. Typical versions of each of these follows:

1. Gravity Model

$$(6.1\text{-}9) \qquad V_{kd}^{e} = V_{k}^{e} \cdot \frac{V_d^e f_{kd}^e}{\sum_d V_d^e f_{kd}^e}$$

where f_{kd}^{e} is an arbitrary function of travel time.

Example

The following table shows typical values of f_{kd}^{e} for a range of travel time, and for three trip purposes: home-based work, home-based nonwork and nonhome-based.

Note that the units of f_{kd}^{e} need not be defined, since the quantities appear in both the numerator and denominator of the gravity model equation. For the same reason, the values can all be scaled up or down by the same constant without affecting the model's results.

2. Opportunity Model

$$(6.1\text{-}10) \qquad V_{kd}^{e} = V_{k}^{e} \exp(-L_e S_d^e) \, [1-\exp -L_e V_d^e]$$

where:

$$S_d^e = \sum_{j} V_j^e = \text{"subtended volume"}$$

j = all

destinations for which $t_{kj} < t_{kd}$

Table 6.1.1

Typical Gravity Model

Travel Time Factors

(Source: Federal Highway Administration)

f^{e}_{kd} :

t_{kd} (minutes)	Work	Non-Work	Non-Home-Based
3	185	220	210
4	150	160	120
5	125	130	100
6	110	90	80
7	100	85	70
8	85	70	60
9	79	60	55
10	67	50	44
11	61	39	38
12	57	35	32
13	50	27	30
14	48	25	26
15	45	21	23
16	10	16	14
17	2	--	5
18	--	--	--

Example:

In the Chicago Area Transportation Study, three market segments were used:

e = 1: trips from home to work and all trips to the central business district (long residential)

e = 2: trips from work to home and all trips from the central business district (long residential)

e = 3: all other trips (short)

Long residential trips have only nonresidential trip ends as their subtended volumes, and vice versa. Short trips have only short trip ends as their subtended volumes. The model parameter for all trips (e = 1 and 2) is 2.5 X 10^{-6}, that for short trips (e = 3) is 20 X 10^{-6}.

The model can therefore be stated mathematically as follows:

$$V_{kd}^{1} = V_{k}^{i} \exp\,[-2.5 \times 10^{-6} S_{d}^{2}]\ \{1 - \exp\,[-2.5 \times 10^{-6} V_{d}^{2}]\}$$

$$V_{kd}^{2} = V_{k}^{2} \exp\,[-2.5 \times 10^{-6} S_{d}^{1}]\ \{1 - \exp\,[-2.5 \times 10^{-6} V_{d}^{1}]\}$$

$$V_{kd}^{3} = V_{k}^{3} \exp\,[-20 \times 10^{-6}\ S_{d}^{3}]\ \{1 - \exp\,[-20 \times 10^{-6} V_{d}^{3}]\}$$

Each of these distribution models are "share" models; they divide the total trips from k,V_k^e, among all distributions using a fraction which, when summed over all destinations, equals one. Travel time by a single mode, usually highway, is typically the only level-of-service variable used, although in some applications, a "generalized cost" has been used which is a linear combination of travel time, distance, and out-of-pocket costs. The level-of-service variable enters the Opportunity Model in an indirect way

only. It affects the ranking of destinations from each origin, which in turn affects the subtended volumes which enter the model directly.

In some applications of both the gravity and opportunity models, adjustments of the V_d^e's are made after initial application of Equation 6.1-9 or 6.1-10 in an attempt to force the total trips to each destination ($V_{.d}^e = \sum_i V_{id}^e$) to equal the original V_d^e's. This constraint is not guaranteed by the functional form of either distribution model. Following adjustments of the original V_d^e's, the equations are applied again. Iteration through application of the equations and adjustment of the original values continues until a desired level of correspondence between each V_d^e and $V_{.d}^e$ is reached.

6.1.3 Modal Split

The third sequential step in the conventional approach is modal split, the prediction of trips by mode from origin to destination. The independent variables are the trip interchanges resulting from the previous step, plus modal level-of-service variables. Symbolically,

$$V_{kdm}^e = f_{em}(V_{kd}^e, L_{kdq}, S_k, A_d) \tag{6.1-11}$$

where V_{kdm}^e = trips of purpose e from origin k to destination d by mode m

V_{kd}^e = results of the trip distribution step

L_{kdq} = level-of-service variables for all modes q between k and d

S_k = socialeconomic variables of travellers in k

A_d = activity system variables in d

Many approaches have been used to develop functional forms, f_{em}, for modal split models. The most commonly used prior to the last three or four years were regression or table look-up models based on the relative levels of service offered by each mode (e.g., reference [4]). Typically, origin zones have been classified by income level and auto ownership, and for each subgroup linear equations or tables are developed which relate fraction of trips by auto and transit to time and cost ratios or differences.

Example

Figure 6.1-2 shows a set of graphs of modal split relationships for work trips based on relative levels of service and on cross-classifications based on the economic status of the traveller and the relative levels of excess travel time. The specific variables used in the figure are:

$$\text{Vertical axis} = \frac{\text{trips by transit}}{\text{trips by all modes}}$$

$$TTR = \frac{\text{total travel time by transit}}{\text{total travel time by auto}}$$

$$L = \frac{\text{excess travel time by transit}}{\text{excess travel time by auto}}$$

L_1 = 0 to 1.5
L_2 = 1.5 to 3.5
L_3 = 3.5 to 5.5
L_4 = 5.5 and over

$$CR = \frac{\text{out-of-pocket cost by transit}}{\text{out-of-pocket cost by auto}}$$

CR_1 = 0.0 to 0.5
CR_2 = 0.5 to 1.0
CR_3 = 1.0 to 1.5
CR_4 = 1.5 and over

EC = economic status (median worker income, 1963 dollars)

EC_1 = \$0 to \$3,100 per year
EC_2 = \$3,100 to \$4,700 per year
EC_3 = \$4,700 to \$6,200 per year
EC_4 = \$6,200 to \$7,500 per year
EC_5 = \$7,500 per year and over

More recently, the following functional form (termed the binary choice logit form) has been used for f_{em}, applied to modal split between auto (a) and transit (t).

(6.1-12) $$P^e_{kdm} = \frac{1}{1 + \exp\ [h_m(L_{kdg})]}$$

and $$h_m(L_{kdg}) = C_m + \Sigma_\ell a^\ell_m (t^\ell_{kdt} - t^\ell_{kda}) + \Sigma_\ell b^\ell_m (c^\ell_{kdt} - c^\ell_{kda})$$

Again, times and costs have been divided into various variables. The constant, C_m, as well as parameters a^ℓ_m and b^ℓ_m, allow the relative characteristics of modes not measured by times and costs (such as comfort and convenience and modal "image") to be represented in the model. The function h_m can be interpreted as a difference in consumer utility between travel by transit and travel by auto.

6.1.4 Route Choice

The final step in the conventional approach is traffic assignment, the prediction of trips by route and thus by link. Symbolically,

(6.5-13) $$V_{ma} = f(V^e_{kdm}, L_{ma})$$

where V_{ma} = volume of traffic of mode m on link a

L_{ma} = level of service on link a for mode m

This equation is deceptively simple; actually, bridging the gap between predictions of O-D volumes (V^e_{kdm}) to link volumes (V_{ma}) represents a significant problem. This is especially true if an attempt is made to take into account the variation of link levels of service (usually simply travel time) with link volumes. When this is not done, the assignment is termed an all-or-nothing assignment without capacity restraint. Various methods of adjusting travel times (alternatively, applying capacity restraint) have been developed. Each one involves either adjusting travel times after a portion of trips are assigned, and then continuing assignment; or averaging a number of complete assignments, each one based on the travel times corresponding to the previous assignment. In present assignment methods, one of the following assumptions is made about how travellers chose a path from origin to destination:

1. Each traveller chooses the minimum path. This assumption leads to a situation where each path used between an O-D pair has equal travel time and all paths not used have a higher travel time.
2. Travellers distribute themselves in such a way that the time and volume on each path between an O-D pair are related as follows:

$$V_{kdr} = \frac{f(t'_{kdr})}{\sum_r f\,(t_{kdr})} \cdot V_{kd} \qquad (6.1\text{-}14)$$

where:

V_{kdr} = volume of flow for O-D pair k-d on route r

t_{kdr} = travel time on route r

$f(t_{kdr})$ = decreasing function of route travel time.

6.1.5 Critical Appraisal of the Traditional Four-Step Models

The approach described above is the most widely used transportation systems analysis approach--the Urban Transportation Model System (UTMS). It has been applied in over 200 cities in the United States and in many other cities around the world. The development and institutionalization of this approach over the last fifteen years is a major accomplishment; it is the first large-scale application of modern systems analysis techniques to problems of the civil sector.

It is useful to examine this approach critically, from the perspective of the equilibrium theory presented in section 3 and the challenge of today's urban transportation problems.

As we have seen in this traditional approach, the travel demand models are structured into a sequence of four submodels called: trip generation, distribution, modal split, and assignment. Essentially, this amounts to estimating V_{kdmr} in a series of "successive approximations": first, V_k, then V_{kd}, then V_{kdm}, and finally V_{kdmr}.

It seems obvious that the following conditions should be met by any such set of demand models and submodels and the corresponding equilibrium calculating procedure:

1. Level of service, L, should enter into every step, including trip generation (unless an analysis of the data indicates in a specific situation that trip generation is, in fact, independent of level of service for all market segments over the full range of levels of service to be studied).
2. The level-of-service attributes used should be as complete as necessary to predict adequately traveller behavior. For example, time reliability, number of transfers, privacy, etc., should be included if empirical evidence indicates these are important.
3. The same attributes of service level should influence each step (unless the data indicates otherwise). For example, transit fares, auto parking charges, walking distances, and service frequencies should influence not only modal split but also assignment, generation, and distribution.
4. The process should calculate a valid "equilibrium" of supply and demand; the same values of each of the level-of-service variables should influence each step. For example, the travel times that are used as inputs for modal split, distribution, and even generation, should be the same as those which are output as results from assignment. If necessary, iteration from assignment back to generation, distribution, etc., should be done to get this equilibrium.
5. The levels of service of every mode should influence demand. Congestion on highway or transit networks, limited capacity (e.g., parking lots), fares, etc., of each mode should (in

general) affect not only its own demand but also the demand for other modes, at all steps (generation, distribution, modal split, and assignment). That is, there should be provision for explicit cross-elasticities.

6. The estimation procedures should be statistically valid and reproducible.

Careful examination of the traditional approach indicates it violates each of these conditions. As a consequence, serious questions can be raised about the biases and limitations of the flow predictions resulting from use of the models in their traditional forms. Nevertheless, this set of models--the UTMS--is still in widespread "production" use in almost all urban transportation planning activities in the United States.

6.2 Group IV: Disaggregate Simultaneous Models

Modelling at the aggregate level uses data for entire zones whose sizes range from fractions of square miles for urban applications to entire metropolitan areas for intercity applications. Modelling at either of these levels of aggregation smoothes out most of the variations of the behavior of the individuals who actually make the travel decisions being modelled. For this reason, much recent demand modelling effort has devloped disaggregate models for predicting travel decisions of individual travellers. Initially, these studies were concerned only with the mode choice decision. The models developed were individual traveller applications of the form shown in section 4.3. When applied to individuals, the dependent variable can only take on the values 0 or 1, requiring a different set of estimation

procedures to be used than with aggregate models. In the initial models of this type, only two modes were included, leading to a binary choice situation. The binary choice logit functional form has been used for these models. More recently, multiple choice models have been developed, using the multiple-choice logit functional form (Ben-Akiva, 1973; McFadden), or multinomial logit, as it is sometimes called.

One feature of these models which is much discussed in the literature is their "independence of irrelevant alternatives" property. This property has two consequences of concern to transportation analysts:

1. If a new alternative (mode, destination, etc.) is added, the percentage decreases in the usage of all existing alternatives will be constant.
2. If one existing alternative is improved, the percentage decrease in the usage of the remaining alternatives will be constant.

The multinomial logit model (MNL) is of this form:

$$p(i:S) = \frac{e^{U_i}}{\sum_{j \varepsilon S} e^{U_j}}$$

where:

x = level of service attributes

0 = Parameters

$U_i = \sum_k O_k X_{ik}$

In the case of travel choices:

$i = (f,d,m,r)$ and

S = FDMR, the full set of all possible travel choices.

Thus, the simultaneous choice MNL becomes:

$$p(f,d,m,r\text{: FDMR}) = \frac{e^{U_{fdmr}}}{\sum\limits_{F,D,M,R} e^{U_{fdmr}}}$$

We can relate this to the sequential disaggregate form by basic probability theory:

$$p(f,d,m,r) = p(f)\ p(d|f)\ p(m|d,f)\ p(r|m,d,f)$$

$$= p_f(X,\theta) \cdot p_d(X,\theta) \cdot p_m(X,\theta) \cdot p_r(X,\theta)$$

the specific choice functions

(p_f, p_d, p_m, p_r) can be shown to be <u>also multinomial logit</u> in form [Ben-Akiva, 1973]

$$p(f:F) = \frac{e^{U_f}}{\sum\limits_{f \epsilon F} e^{U_f}}$$

$$p(d:D) = \frac{e^{U_{df}}}{\sum\limits_{d \epsilon D} e^{U_{df}}}$$

$$p(m:M) = \frac{e^{U_{mdf}}}{\sum\limits_{m \epsilon M} e^{U_{mdf}}}$$

$$p(r:R) = \frac{e^{U_{rmdf}}}{\sum\limits_{r \epsilon R} e^{U_{rmdf}}}$$

The first estimation of a simultaneous disaggregate model was by Ben-Akiva (1973). The functional form of this multiple-choice logit model is as follows:

$$(6.8\text{-}1)\quad \frac{P_{kdm}}{P_{kd'm'}} = \exp[\sum_q a\ (A_{dq} - A_{d'q}) + \sum_q b_q\ (M^1_{mq} - M^1_{m'q})$$

$$+ \sum_q c_q Y_k (M^2_{mq} - M^2_{m'q}) + \sum_q h_q (L^1_{kdmq} - L^1_{kd'm'q}) + \sum_q \frac{fg}{Y_k}\ (L^2{}_{kdmq} - L^2{}_{kd'm'q})]$$

where:

P_{kdm}, $P_{kd'm'}$ = fraction of total trips from household k going to destinations d and d' by modes m and m'. (Either d abd d' or m and m' may be the same, but not both.)

A_{dq}, $A_{d'q}$ = activity system variables

M^i_{mq}, $M_{m'q}$ = modal variables

L^i_{kdmq}, $L^i_{kd'm'q}$ = level of service variables

Y_k = household income code 9

As estimated by Ben Akiva, the following variables were used:

1. Activity system variables (A_{dq}):

   A_{d1} = number of jobs in wholesale and retail establishments in the zone of destination d

   A_{d2} = indicator for DBD destinations. (1 if d = CBD, 0 otherwise)

2. Modal variable in separate term (M^1_{mq}):

   M^1_{m1} = indicator for auto usage. (1 if m = auto, 0 otherwise)

3. Modal variable in interaction term with income (M^2_{mq}):

   M^2_{m1} = indicator for auto usage. (same as M^1_{m1})

4. Level of service variables in separate terms (L^1_{kdmq}):

L^1_{kdm1} = out-of-vehicle travel time

L^1_{kdm2} = in-vehicle travel time.

5. Level of service variable in interaction term (L^2_{kdmq}):

L^2_{kdm1} = out-of-pocket cost.

This model was estimated for auto and transit trips for the shopping purpose only, and does not deal with trip-making frequencies or time of day choices. It therefore represents a model which can be used to divide total shopping trips from a household among the variable modes and destinations. The parameters obtained are the following:

Associated Variable	Parameter Level	Parameter Value
$(A_{d1} - A_{d'1})$	a_1	.000171
$(A_{d2} - A_{d'2})$	a_2	.316
$(M^1_{m1} - M^1_{m'1}$	b_1	-1.36
$Y_k(M^2_{m1} - M^2_{m1})$	c_1	.114
$(L^1_{kdm1} - L^1_{kd'm'1})$	h_1	-.0633
$(L^2_{kdm2} - L^2_{kd'm'2})$	h_2	-.0164
$L^2_{kdm1} - L^2_{kd'm'1})/Y$	f_1	-.0757

In additional work since that time, Ben-Akiva and his associates have developed additional simultaneous choice disaggregate models. Some of these models will be summarized briefly here and are described in detail in the appendix.

The three models described below are:

1. A work mode choice model with an explicit carpool mode
2. A simultaneous choice model for shopping trips including the choices of travel frequency, trip destination, and mode of travel
3. A household auto ownership model

All these models were estimated using the multinomial logit model with data sets derived from the 1968 Washington, D.C. Home Interview Travel Survey, conducted by the Metropolitan Washington Council of Governments. In addition, in recent work, a model of simultaneous choice of residential location, auto ownership, and mode to work has been developed (Lerman).

1. Work Choice Model[10]

A work mode choice model has been produced as a multinomial-logit three-mode (auto driver alone, carpool, transit) model, estimated using the Washington, D.C. data base.

The model specification recommended is shown in Appendix A. This model contains all the normally expected variables--in-vehicle travel time, out-of-vehicle travel time, out-of-pocket costs, income, and auto availability--plus some special variables to pick up some of the factors influencing carpooling. Further description of this model and a discussion of the variables can be found in Appendix A.

It should be noted that, since the models produce choice probabilities, they do not assume that a traveller will use the same mode every day. This is particularly important for carpool users in view of the fact that many existing carpools have been observed

not to be an everyday arrangement and carpools are frequently organized for less than five trips per week. Thus, if a traveller carpools twice a week, for example, his carpool choice probability on any given day is 0.4.

2. Non-Work Trip Frequency, Destination, and Mode Choice Models

Two nonwork (shopping) simultaneous disaggregate travel demand models have been developed on the 1968 Washington, D.C. data base. The first was the estimation of a joint destination and mode choice shopping model by Ben-Akiva (1973), described above. The second was the subsequent estimation of a joint frequency, destination, and mode choice nonwork model by Adler and Ben-Akiva (1974). Both are logit models calibrated by maximum likelihood estimation and both forecast the frequency, destination, and mode choice of a household's shopping trip as a function of automobiles available, among other factors. The second model was estimated on a larger number of observations than the first, has reasonable coefficients in terms of both sign and relative magnitude, and all the important coefficients are statistically significant.

The variables used in this second model and the estimation results are shown in tables B.1 and B.2 in Appendix B. The model contains transportation level-of-service variables, socioeconomic characteristics and shopping attraction variables.

Particularly noteworthy in the specification of this model are its relationships with work trips. The auto availability variable is formulated as the auto available to the household minus those used for work trips by household member. This specification assumes

that the choice of mode with respect to work trips is of higher importance to the household than the shopping travel choices. This specification allows for an explicit evaluation of a policy that discourages the use of car for peak-hours trips and thereby increases the availability of auto for off-peak nonwork trips. This model would be used after the work mode split model when auto availability was determined so that it could be input to this shopping model. Further description of this model and a discussion of the variables can be found in Appendix B.

3. Auto Ownership Model

Recently, a simultaneous disaggregate auto ownership and work mode choice model has been developed. This model addresses the classic "chicken and egg" dilemma of the close interdependency between auto ownership and mode-to-work choice by developing and estimating a disaggregate choice model which assumes these two decisions are made simultaneously. Each combination of auto ownership level and usual mode to work is a distinct "bundle" or alternative, one and only one of which is assumed to be selected by each household.

The model describes the joint probability of a household selecting a given auto ownership level and a given mode to work for the breadwinner. (The worker with the highest occupation status code has termed the "bread-winner"; it was felt that his transportation to work would dominate in terms of demand for the family auto.) Hence, the choice set for a household consists of the cross-product of the entire set of modes and the entire set of possible

auto ownership levels. However, this can be reasonably simplified without any major loss in the model's usefulness by assuming that a maximum of eight choices forms the set of alternatives which is assumed to be available to any household. These are as follows:

1. Own zero autos and carpool to work
2. Own zero autos and use transit to work
3. Own one auto and drive alone to work
4. Own one auto and carpool to work
5. Own one auto and use transit to work
6. Own two or more autos and drive alone to work
7. Own two or more autos and use transit to work

Note, however, that there is no requirement in the logit model that every household have all eight alternatives. For some work trips, transit service is simply not likely to be perceived as available; these households choose among only five alternatives. Other households have an income which is simply too low for them to consider realistically multiple car ownership as an alternative; these households have only the first five alternatives. A variety of other combinations is also feasible. For example, in the model calibration it was assumed that a household will not own more cars than the number of licensed drivers who might use those cars. These rules for determining which alternatives pertain to each household were verified by performing extensive cross-tabulations on the data set and examining whether any substantial portion of the sample made a choice assumed to be unavailable.

A broad range of variables was considered for providing possible candidates in the model specification. An attempt was made to restrict the model to variables which are behaviorally related to the household decision process rather than simply correlated with actual causal variables.

The model specification recommended is discussed in Appendix C. The important variables contained in the model include the transport level-of-service characteristics (in-vehicle travel time, out-of-vehicle travel time, and travel cost) for both peak and off-peak travel. In addition, the model includes other factors that influence carpooling, and variables that determine auto ownership and mode choice such as housing and socioeconomic characteristics of the household.

One of the most important aspects of the auto ownership model development was the determination of significantly, yet logically different, behavioral responses from different market segments (groups with homogeneous socioeconomic characteristics). The above model was estimated for a pooled sample of households but, in addition, separate models were estimated for each of several different market segments. This same stratification by market segment could be applied to all three models (work mode split, shopping direct demand, and auto ownership) to yield more specific information on response to the various policies.

Further description of the joint auto ownership/work mode choice models and a discussion of the variables can be found in Appendix C.

Note that, while the model is formulated here as a joint (auto ownership and work mode choice simultaneously) model, separate marginal (auto ownership alone or work model choice alone) or conditional (auto ownership, given work mode choice; or work mode choice, given auto ownership) models are directly and easily derivable. Moreover, the work mode choice portion of this joint model refers only to the mode choice decision of the breadwinner, not of all household workers. This is in contrast to the work mode choice model for all household workers. This gives rise to two alternative ways of utilizing these three available models for immediate planning analyses:

1. Use the marginal auto ownership form of the joint model to yield auto ownership; then use the section 1. work mode split model for all workers, yielding complete information on work trip mode splits and auto availability for other trips; then run the nonwork demand model to capture changes in nonwork travel due to increased auto availability.
2. Use the joint auto ownership/breadwinner work mode choice model for breadwinner work mode split; use the section 1. work mode split model for all other workers (since there is a breadwinner variable in that model, there is no bias in using it for nonbreadwinners only): then, taking the composite results of those models in terms of remaining auto availability, run the nonwork demand model.

4. Features of These Models

It is useful to summarize the major advantages of these models over other travel demand models that have been developed:

1. Car Pooling

The work mode choice model includes a car pool mode in addition to drivers driving alone and transit. Other existing work mode choice models consider auto drivers and auto passengers as two separate modes. Therefore, one cannot isolate persons in multiple occupancy vehicles from those with a driver only. Analysis of special incentives to multiple occupancy vehicles can only be performed using a model with an explicit "car pool" mode.

2. Auto Ownership

These models include an auto ownership model sensitive to the important transportation policy variables, while auto ownership is assumed as an exogenous (independent) variable in other travel demand models. The importance of this model for the analysis of today's policy thrusts is obvious: auto ownership can be influenced by transportation policies, and these models include this effect.

3. Sample Size

In a recent project to estimate disaggregate travel demand models in the Netherlands, an analysis of the effect of sample size on the reliability of the estimated coefficients was performed (Ben-Akiva and Richards, 1975). It was found that, for a multiple logit mode choice model with ten

coefficients, samples of less than 250 observations result in very unstable coefficients. The standard errors of the estimated coefficients decreased sharply by increasing the sample size gradually from 100-150 to 250-300 observations. Beyond 300, the incremental reductions in the standard errors with increasing sample size were much smaller and gradually diminishing. Disaggregate models are clearly very attractive from the perspective of data collection economics.

4. Explanatory Power

These models are specified using a greater variety of explanatory behavioral variables than other available models. The improved specifications result in increased explanatory power and reduced uncertainty for forecasting.

5. Theoretical Validity

The assumption of a joint (or simultaneous) decision process, made in the shopping and auto ownership models, is more realistic than the sequential decision process implied by the structure of other available models, both aggregate and disaggregate.

6. Market Segments

The auto ownership model was estimated using nine market segments. The other models also use a richer set of socioeconomic variables than other existing models in the field. This allows for a better understanding of how

various population groups will respond differently to a given policy option.

7. Transferability

Because the coefficients of a disaggregate model are free from aggregation bias, a disaggregate model can potentially be applied to situations (i.e., geographical areas) which are different from those on which it was calibrated. However, even a disaggregate model can be transferred only if it is not seriously misspecified. These disaggregate travel demand models are based on the most complete specifications among all similar models that are known at this time. Therefore, we would expect these models to produce reliable forecasts for other urban areas in addition to Washington, D.C. for which the models were estimated. Some tests that were carried out with these models on data from New Bedford, Mass., Milwaukee, Portland, and Los Angeles show that these models can be used satisfactorily as estimated, or with minor adjustments.[11]

8. Trip Frequency

For the shopping trips, the frequency of travel is not assumed constant as in the UTMS trip generation models.

7. Conclusions

This paper has presented an overview of the major choices for development of travel forecasting methods:

1. In the past, the basic approach used for urban travel forecasting has been that of sequential, aggregate, deterministic demand models with an indirect approach to calculating equilibrium flows. This approach now is seen to have serious deficiencies, especially when confronted with today's emphasis on a broad range of transportation options.
2. Recent research indicates that there are at least eight alternative ways to develop demand models and use them for equilibrium prediction.
3. From the point of view of equilibration, a simple-step direct approach is likely to be most valid and most efficient in all except special situations.
4. From the point of view of demand model development, estimation of models using disaggregate methods is much more economical in terms of data requirements than aggregate methods, and provides the opportunity for formulation models soundly based on behavioral theory and estimated using valid statistical approaches.
5. Numerous practical approaches to aggregation of disaggregate models to produce an aggregated model for forecasting are available.
6. It is feasible to estimate simultaneous choice disaggregate models, and the results are different from the results of separately estimated sequential models. The assumption of simultaneous choice is the soundest behavioral assumption except when there are specific behavioral reasons for

hypothesizing a particular sequence of choices. For example, in current work, one hypothesis being used is of a simultaneous-block-sequential form:

A. simultaneous choice of residential location, employment location, and housing type

B. simultaneous choice of automobile ownership level and mode to work, sequential to (conditional on) the choices in (A)

C. simultaneous choice of non-work travel, including frequency, destination, mode, time of day, and route of nonwork trips, sequential to (conditional on) the prior choices in (A) and (B).

Alternative behavioral hypotheses might be appropriate in different contexts.

As a consequence of the above, the author's conclusion is: The preferred approach, in general is to:

1. Formulate a simultaneous or simultaneous-block-sequential form based upon behavioral reasoning in a particular context
2. Estimate the corresponding models in disaggregate simultaneous (and sequential, where appropriate) forms
3. Construct models for prediction by aggregating explicitly the estimated disaggregated models
4. Predict flow impacts of alternative transportation options by equilibrating the resulting aggregate simultaneous models in a single-step direct approach.

Where specific situations suggest that short cuts may be acceptable, use appropriate approximation methods. For example, proposed changes in transit service frequency in a particular corridor may be hypothesized as unlikely to affect choices of destination or auto ownership level; therefore, a conditional mode choice model could be derived from the simultaneous models and used in a sequential approach to equilibrium with trip generation and distribution held fixed. However, for other changes, this hypothesis might be considered unacceptable, and so the more general approach should be followed.

APPENDIX A

Work Mode Split Model

Tables A.1 and A.2 contain the model specification (variables included) and estimation results.

Interesting results were obtained for the implied impact of changing auto availability and also for the importance of the "government employee" variable. Both show considerable influence on the modal split forecast by the model, and both have strong (significant) coefficients. GW is a dummy variable in the car pool utility function which is unity if the traveller is a civilian employee of the federal government. Its positive sign indicates that, all else constant, federal employees are more likely to choose carpooling than nonfederal employees, which may be due to more potential for matching (i.e., more opportunities to choose from and less search time), more active encouragement of carpooling by the employer, or more consistent start-work and end-work times. An important implication of the strength of this variable is that concerted efforts by large employers to encourage and facilitate carpooling can have a significant impact on the proportion of work trips made by that mode. This variable, in effect, is a proxy reflecting the existence of some employer-based carpool incentive programs. Therefore, in forecasting, this variable can be used as a policy variable to indicate those employers who instituted such programs by setting its value equal to one (or, if sufficient empirical data exists, to some other value reflecting the level of the employer's incentive program relative to the federal government's).

DEFINITION OF VARIABLES

Variable Code	Definition
1. D_c	= 1, for drive alone 0, otherwise
2. D_s	= 1, for shared ride 0, otherwise
3. OPTC/INC	= round trip out-of-pocket travel cost (in cents)/household annual income (in dollars)
4. IVTT	= round trip in-vehicle travel time (in minutes)
5. OVTT/DIST	= round trip out-of-vehicle travel time (in minutes)/one way distance (in miles)
6. $AALD_c$	= # of autos/licensed drivers, for drive alone 0, otherwise
7. $AALD_s$	= # of autos/licensed drivers, for shared ride 0, otherwise
8. BW_c	= 1, if worker is the breadwinner, for drive alone 0, otherwise
9. GW_s	= 1, if worker is a civilian employee of the federal government, for shared ride 0, otherwise
10. $DCITY_c$	= 1, if work place is in the CBD, for drive alone 0, otherwise
11. $DCITY_s$	= 1, if work place is in the CBD, for shared ride 0, otherwise
12. $DINC_{c,s}$	= household annual income - 800 * # of persons in the household (in $), for drive alone and shared ride 0, otherwise
13. $NWORK_s$	= # of workers in the household, for shared ride 0, otherwise
14. $DTECA_s$	= employment density at the work zone (employees per commercial acre) * one way distance (in miles), for shared ride 0, otherwise

Alternative:

c = drive alone
s = shared ride (car pool)
t = transit

WORK MODE CHOICE MODEL

THE MODEL COEFFICIENTS

Variable		Coefficient	t-statistic
1. Drive alone constant	D_c	-3.24	-6.86
2. Shared ride constant	D_s	-2.24	-5.60
3. Out-of-pocket travel cost divided by income	OPTC/INC	-28.8	-2.26
4. In-vehicle travel time	IVTT	-.0154	-2.67
5. Out-of-vehicle travel time divided by distance	OVTT/DIST	-.160	-4.08
6. Auto availability (drive alone only)	$AALD_c$	3.99	10.08
7. Auto availability (shared ride only)	$AALD_s$	1.62	5.31
8. Breadwinner (drive alone only)	BW_c	.890	4.79
9. Government worker (shared ride only)	GW_s	.287	1.78
10. CBD work place (drive alone only)	$DCITY_c$	-.854	-2.75
11. CBD work place (shared ride only)	$DCITY_s$	-.404	-1.36
12. Disposable income (drive alone and shared ride only)	$DINC_{c,s}$	.0000706	3.46
13. Number of Workers (shared ride only)	$NWORK_s$	.0983	1.03
14. Employment density (shared ride only)	$DTECA_s$	.000653	1.34

. of observations = 1114
# of alternatives = 2924

The auto availability variable occurs twice (with different coefficients) because of its mode specific character, once in the drive-alone utility function, and once in the carpool utility function. Both coefficients have positive signs and, therefore, indicate that an increase in the ratio of autos to drivers in a household will increase the likelihood of both driving alone and in a carpool, relative to using public transit. However, because the difference between the coefficients for driving alone and for carpooling is also positive, an increase in auto availability (all else equal) will raise the probability of driving alone relative to that of participating in a carpool.

The level of service variables: out-of-pocket travel cost (OPTC), in-vehicle travel time (IVTT), and out-of-vehicle travel time (OVTT) have significant coefficients with the expected negative signs and relative magnitudes. For a typical five-mile commuting trip, the ratio of the out-of-vehicle travel time coefficient to that of in-vehicle time is approximately 2. This implies that a minute of out-of-vehicle travel time is twice as onerous as a minute of in-vehicle time. This result is similar to results obtained in previous studies. The implied value of time for a five-mile commuting trip by an individual whose annual household income is $8000 (in 1968 dollars) is approximately $2.50/hr. for in-vehicle time (and, of course, twice that for out-of-vehicle time). Again, these figures correspond to logic and the results of previous mode choice models.

The breadwinning variable is included for drive alone with a positive coefficient, which represents the higher priority of the primary worker for the use of any car that is available.

The CBD variable captures the added disutility of driving to the CBD which is quite likely to differ between the two auto modes, but nevertheless captures the disutility of driving in the CBD. As expected, the coefficient of the variable for drive alone is more negative than that for carpool.

Disposable income is included with one coefficient because when separate coefficients were tried they were very similar. The number of workers has been included to capture the effects of intrafamily carpools. The final variable is an interaction variable that combines the added convenience of finding a carpool partner if one works in dense employment areas with the added incentive to carpool over longer distances. This variable could be improved by collecting more data on employer characteristics which were not available in the existing data set.

The two modal constants, for drive alone (D_c) and for shared ride (D_s) are negative. The larger absolute value of the drive alone pure (dis)preference indicates it is relatively less desirable than carpooling, given that all other variables are constant across all modes. Both drive alone and carpool are less desirable than transit (implied zero constant) as far as the constant term alone goes, but the auto availability variables pick up enough positive value when an auto is available to reverse this relative effect for drive alone and render carpool only slightly less desirable than transit.

APPENDIX B

Shopping Trip Frequency, Destination, and Mode Choice Model

The model specification and the estimation results are shown in Tables B.1 and B.2.

As can be seen, all the important variables are significant at the 99% confidence interval (t-statistic greater than 2.3). The co-efficients of the level-of-service variables (time, cost) have the expected negative signs and result in reasonable values of time. For a shopping trip of 2.5 miles and a total round trip travel time of 40 minutes, a minute of out-of-vehicle time is somewhat less than twice as onerous as a minute of in-vehicle time.

The attraction variable, ln (REMP), has the correct sign (the higher the retail employment at a given shopping center, the more attractive it is). The car constant, DC, has a negative coefficient, reflecting that when the auto is used for a shopping trip it is not available for other purposes. The CBD constant is positive, indicating that the CBD is a more attractive shopping area generally. The AAC variable coefficient is positive, meaning that the greater the number of autos available, the more likely the traveller is to make a shopping trip and make it by auto. The positive sign on the zero-frequency income variable INCF, is consistent with other results and states that higher income households make fewer shopping trips (i.e., they are able to maintain larger stocks of goods).

Of the other frequency variables, household size (HHSF) enters with a negative sign indicating (as expected) that, for a larger household, the probability of not making a shopping trip on a given day is less. The density variable DENF is an attempt to account for

Table B.1

SHOPPING JOINT TRAVEL DEMAND MODEL DEFINITION OF VARIABLES

Variable Code		Definition
1. DC	=	1, for car 0, otherwise
2. OVTT/DIST	=	round trip out-of-vehicle travel time (in minutes)/one way distance (in miles)
3. IVTT+OVTT	=	round trip in-vehicle travel time + round trip out-of-vehicle travel time (in minutes)
4. OPTC/INC	=	round trip out-of-pocket travel cost (in cents)/household annual income (in code*)
5. AAC	=	# of autos available to household - # of autos used for work trips by workers in the household, for car 0, otherwise
6. 1/DIST	=	1/one way distance (in miles)
7. REMP	=	retail employment at shopping destination (in # of employees)
8. DCBD	=	1, for CBD shopping destination 0, otherwise
9. DF	=	1, for zero frequency 0, otherwise
10. HHSF	=	# of persons in the household, for zero frequency 0, otherwise
11. DENF	=	retail employment density in residence zone (employees per acre), for zero frequency 0, otherwise
12. INCF	=	household annual income (in code*), for zero frequency 0, otherwise

Alternatives:

No trip (zero frequency), variables (1) - (8) are equal to zero. Trip to shopping destination d and by mode m, for all relevant shopping destinations (including the CBD), and for car and transit modes.

* Income code (in 1968 $)

1 =	0 - 2,999	6 =	10,000 - 11,999
2 =	3,000 - 3,999	7 =	12,000 - 14,999
3 =	4,000 - 5,999	8 =	15,000 - 19,999
4 =	6,000 - 7,999	9 =	20,000 - 24,999
5 =	8,000 - 9,999	10 =	25,000 +

Table B.2

SHOPPING JOINT TRAVEL DEMAND MODEL

THE MODEL COEFFICIENTS

Variable		Coefficient	t-statistic
1. Car constant	DC	-.555	-2.13
2. Out-of-vehicle travel time divided by distance	OVTT/DIST	-.100	-3.38
3. Total travel time	ln (IVTT+OVTT)	-2.24	-11.85
4. Out-of-pocket travel cost divided by income	OPTC/INC	-.0242	-4.20
5. Autos available for non work trips (car only)	AAC	.557	5.61
6. Inverse of distance to destination	1/DIST	6.86	1.66
7. Retail employment at destination	ln (REMP)	.161	3.29
8. CBD destination constant	DCBD	.562	2.07
9. Frequency zero constant	DF	-3.78	-4.51
10. Household size (frequency zero only)	HHSF	-.186	-4.57
11. Retail employment density at the origin (frequency zero only)	DENF	.000598	1.38
12. Household income (frequency zero only)	INCF	.0414	1.18

# of observations = 1313

# of alternatives = 44718

walk trips to shopping, which are not recorded in the Home Interview Survey. This variable measures the density of retail employment in the home zone, and thus is a proxy for the availability of suitable shopping destinations within walking distance of the home. The positive sign of the DENF coefficient means that a household living in a zone with dense retail employment is less likely to embark on a vehicular shop trip (i.e., is more likely to choose a walk shop trip instead).

APPENDIX C

Auto Ownership Model

Table C.1 lists the variables included in the model. The subscripts for each variable indicate the alternatives for which the variable enters into the utility function. Variables with no subscripts enter into every utility. Seven constants (one for each alternative except the zero autos-transit to work alternative) measure a constant relative effect which the remaining variables do not pick up. These variables represent what might be termed a "pure alternative effect."

The variables $AALD_c$ and $AALD_s$ measure the number of autos per licensed driver, thereby modeling the auto availability in the mode choice decision. The more autos per driver a household has, the greater the probability the primary worker will use a car to go to work; hence, a positive sign would be anticipated.

The variable denoted as Z requires some explanation. It represents the discretionary income left to a household after transportation and fixed household expenditures. The rationale underlying the use of this variable is quite simple. The total utility a household derives from any auto ownership-mode choice selection depends on the attributes of the alternative, <u>one of which is the resources remaining to the household for other uses, if that alternative is selected</u>. The use of the Z variable is based on the theory that the marginal utility of money at a given point on the income curve is the same for all categories of household expenditures. While it is theoretically possible to incorporate a fuller set of alternatives and thereby eliminate

Table C.1

AUTO OWNERSHIP MODEL

DEFINITION OF VARIABLES

Variable Code		Definition
1. D_{0s}	=	1, for zero autos; shared ride 0, otherwise
2. D_{1c}	=	1, for one auto; drive alone 0, otherwise
3. D_{1s}	=	1, for one auto; shared ride 0, otherwise
4. D_{1t}	=	1, for one auto; transit 0, otherwise
5. D_{2c}	=	1, for two autos; drive alone 0, otherwise
6. D_{2s}	=	1, for two autos; shared ride 0, otherwise
7. D_{2t}	=	1, for two autos; transit 0, otherwise
8. $AALD_c$	=	# of autos/licensed drivers, for drive alone 0, otherwise
9. $AALD_s$	=	# of autos/licensed drivers, for shared ride 0, otherwise
0. Z	=	household annual income - 800 * # of persons in the household - 1000 * # of autos - 250 * daily round trip travel cost (in $)
1. HT_2	=	1, if household lives in a single family house, for two autos 0, otherwise
2. IVTT	=	daily round trip in-vehicle travel time (in minutes)
3. OVTT/DIST	=	daily round trip out-of-vehicle travel time (in minutes)/one way distance (in miles)
4. AALD	=	# of autos/licensed drivers
5. R_1	=	car generalized shopping travel cost/transit generalized shopping travel cost, for one auto 0, otherwise

Table C.1(continued)

Variable Code		Definition
6. R_2	=	car generalized shopping travel cost/transit generalized shopping travel cost, for two autos 0, otherwise
7. $DCITY_c$	=	1, if work place is in the CBD, for drive alone 0, otherwise
8. $DCITY_s$	=	1, if work place is in the CBD, for shared ride 0, otherwise
9. TOPTC	=	250 * daily round trip out-of-pocket travel cost (in $)
0. GW_s	=	1, if worker is a civilian employee of the federal government, for shared ride 0, otherwise
1. $NWORK_s$	=	# of workers in the household, for shared ride 0, otherwise
2. $DTECA_s$	=	employment density at the work zone (employees per commercial acre) * one way distance (in miles), for shared ride 0, otherwise

Alternative		
0s	=	zero autos; shared ride to work
0t	=	zero autos; transit to work
1c	=	one auto; drive alone to work
1s	=	one auto; shared ride to work
1t	=	one auto; transit to work
2c	=	two + autos; drive alone to work
2s	=	two + autos; shared ride to work
2t	=	two + autos; transit to work

possible expenditures on commodities not explicitly modelled, it is undesirable to do so in a practical model.[12] Hence, some relatively simple means of collapsing the entire set of ignored consumption choices is required.

This is precisely what the Z variable does. It incorporates other choices into a term which represents the income remaining to a household after certain fixed household and transportation expenditures. It thus provides a way of including transportation costs versus capital and operating--directly in the model.

An alternative with zero auto ownership results in a high value of Z, reflecting the availability of income the household would have to allocate to the purchase, maintenance, and operation of a car if it had chosen to do so. Thus, the coefficient of the Z variable in the utility function should have a positive sign.

The value of the Z variable was also used to evaluate whether or not a household had sufficient income to preceive an alternative as being available. Any alternative for which Z was less than zero was excluded from the household's choice set.

HT_2 is a 0,1 dummy variable representing single family residence. It was hypothesized that single family residence reduces parking difficulties and increases the need for multiple automobile ownership for social and recreational trips. The dummy variable only applied for multiple car alternatives.

In-vehicle travel time, out-of-vehicle travel time, and out-of-pocket travel cost are self-explanatory. These variables enter the model with the same coefficients for the two modes. This is based on the results obtained by ORA (1972) and Ben-Akiva and Richards (1975).

The variable AALD is the number of autos divided by the number of licensed drivers in the household and measures the need for additional autos in the household. As the number of autos per driver decreases, the need for multiple car ownership should increase. The difference between AALD and the mode specific $AALD_c$ and $AALD_s$ variables is that they appear in different equations as measures of different effects, the former an incentive for owning more cars and the latter an auto availability effect for mode choice. One would anticipate that these variables are highly colinear in the data set and that the estimates of their coefficients would have high standard errors. However, this did not turn out to be the case. The reason for this is that the variables are defined to be zero for different alternatives. $AALD_c$ is set at zero for alternatives (1) and (2) only. The modal choice effect $AALD_c$ and $AALD_s$ indicated increased car mode preference when autos per licensed driver increased, while the auto ownership effect (AALD) indicated increased auto ownership preference when number of licensed drivers increased. The estimates of the coefficients of AALD and $AALD_c$ were highly significant.

The next two variables, R_1 and R_2, incorporate the effects of shopping trip accessibility and thus represent nonwork travel opportunities. These variables were developed by forecasting the attributes of expected (or average) shopping trip for each household by both car and transit. This was done using a previously estimated simultaneous shopping destination and mode choice model estimated by Ben-Akiva (1973). The attributes of these expected trips (travel time, cost, etc.) were then weighted using the coefficients of the previously estimated model. Hence, the ratio of car cost to transit

cost was taken in order to measure the relative accessibility of car shopping travel to that of transit. Note that when transit is not available to a household for shopping travel, the ratio is zero (i.e., infinite transit cost).

Expressed mathematically, the expected value of attribute i for a given mode m, X_{im}, is as follows:

$$E[X_{im}] = \sum_{d \in D} P(d|m)\, X_{idm},$$

where D is the set of shopping destinations which can be reached by mode m and P(d|m) is the probability of making a shopping trip to destination d given that mode m is used. The generalized price by mode m, GP_m, is therefore the following:

$$GP_m = \alpha_1\, E[IVTT_m] + \alpha_2\, E[OVTT_m] + \alpha_3\, E[OPTC_m]$$

where $E[IVTT_m]$ is the expected shopping in-vehicle travel time by mode m;

$E[OVTT_m]$ is the expected shopping out-of-vehicle travel time by mode m;

$E[OPTC_m]$ is the expected shopping out-of-pocket travel cost by mode m; and

$\alpha_1, \alpha_2, \alpha_3$ are the parameters of $IVTT_m$, $OVTT_m$ and $OPTC_m$, respectively, from a previously estimated shopping travel demand model.

Finally, the ratio, R, is simply defined as

$$R = \frac{GP_{car}}{GP_{transit}}$$

The concept of generalized prices was used by both CRA (1972) and Ben-Akiva (1973).

The CBD dummy variables are used to reflect the disutility typically associated with downtown car usage which is not entirely measured by travel times and cost. Effects such as high travel time variance, the frustration drivers experience in congestion, and the unreliability of on-street parking are all incorporated into these variables.

NOTES

1. M. Beckman, C. B. McGuire, and C. B. Winston, Studies in the Economics of Transportation, New Haven: Yale University Press (1956); G. Kraft and M. Wohl, "New Direction for Passenger Demand Analysis and Forecasting," Transportation Research Vol. 1 (1967), pp. 205-230; M. Wohl and B. V. Martin, Traffic Systems Analysis, New York: McGraw-Hill (1967), Chapter 5; M. L. Manheim, "Search and Choice in Transport Systems Planning," Highway Research Record 293, Washington, D.C.: Highway Research Board (1969).

2. Manheim, op. cit.

3. We deal here with only one of "Wardrop's Principles"; we will not discuss the other, concerned with global optimization of the flow pattern, as this is inapplicable to urban transportation flow prediction. Cf. J. G. Wardrop, "Some Theoretical Aspects of Road Traffic Research," Institute of Civil Engineers, Road Paper No. 36, London (1952); M. Beckmann, "On the Theory of Traffic Flow in Networks," Traffic Quarterly (January, 1967); W. S. Jewell, "Models for Traffic Assignment," Transportation Research Vol. 1, No. 1 (May, 1967).

4. B. V. Martin, F. W. Memmott and A. J. Bone, Principles and Techniques of Predicting Future Urban Area Transportation, Cambridge, Mass.: M.I.T. Press (1965); various publications of the U.S. Federal Highway Administration on Urban Transportation Planning.

Notes (continued)

5. In the equations which follow, standard conventions of probability theory are used. The relevant concepts, explained briefly here, are described in any standard text on probability.

   P(A) is the probability that a random event, A, occurs.
   The joint probability P(A,B) is the probability that two random events, A and B, both occur.
   The conditional probability P(A|B) is the probability of event A occuring when it is known that event B occurs.

6. The conditions are that f_1, f_2, f_3, and f_4 in (4.5-3) be "internally consistent" as defined by Manheim (1973).

7. Adapted from Earl R. Ruiter (1973), HRB Special Report 143; Cambridge Systematics, UMTA Travel Forecasting Manual; and Marvin L. Manheim, Fundamentals of Transportation Systems Analysis.

8. Usually per day.

9. The following code was used:

Household Income (in \$1,000/year)	Code
0 - 3	1
3 - 4	2
4 - 6	3
6 - 8	4
8 - 10	5
10 - 12	6
12 - 15	7
15 - 20	8
20 - 25	9
greater than 25	10

Notes (continued)

10. Binary work mode choice models have been developed by many researchers; this model is described here because of its relationship to the simultaneous models described following, and because it is the first disaggregate model to treat carpools as an explicit mode.

11. A variety of procedures to adjust a model to a different area are developed and evaluated in Atherton (1974).

12. A more formal case for the use of the Z variable can be made by utilizing the concept of a utility "tree," in which a household's utility function is considered to be additively separable. See Strotz (1957) and Cambridge Systematics (1974).

REFERENCES

Adler, T. and M. Ben-Akiva. 1974. "A Joint Frequency, Destination and Mode Choice Model for Shopping Trips," paper submitted to the Transportation Research Board.

Aldana, E., Richard deNeufville, and Joseph Stafford. 1973. "Micro-analysis of Urban Transportation Demand." Highway Research Board 1973.

Atherton, Terry. 1974. "A Bayesian Approach to Transferring Stochastic Disaggregate Travel Demand Models," unpublished S.M. thesis, M.I.T.

Beckmann, M.J., C. B. McGuire, C. B. Winston. 1956. Studies in the Economics of Transportation. Yale University Press.

Beckmann, M. J. 1967. "On the Theory of Traffic Flow in Networks." Traffic Quarterly. January.

Ben-Akiva, Moshe E. 1973. "Structure of Passenger Travel Demand Models." Ph.D. dissertation, M.I.T.

Ben-Akiva, Moshe E. and M. G. Richard. 1975. "A Disaggregate Multi-Modal Model for Work Trips in The Netherlands." Transportation Research Board.

Buro Goudappel en Coffeng and Cambridge Systematics, Inc. 1974. Disaggregate and Simultaneous Travel Demand Models: A Dutch Case Study. Cambridge, Mass.

Cambridge Systematics, Inc. 1974. Introduction to Urban Travel Demand Forecasting, a tutorial manual prepared for UMTA by Cambridge Systematics, Inc., March.

Cambridge Systematics, Inc. 1974. Multinomial Logit Estimation Package: Program Documentation. Version 2, Mod 1. Cambridge, Mass.

Cambridge Systematics, Inc. 1974. Series of auto ownership technical memoranda.

Charles River Associates, Inc. (CRA). 1972. A Disaggregated Behavioral Model of Urban Travel Demand. Federal Highway Administration, U.S. Department of Transportation, Washington, D.C.

Charles River Associates, Inc. 1967. A Model of Urban Passenger Travel Demand in the San Francisco Metropolitan Area. Report Number 117-1. Prepared for Peat, Marwick, Livington & Co., December.

Dafermos, S. C. 1972. "The Traffic Assignment Problem for Multi-class User Transportation Networks," Transportation Science. vol. 6, No. 1, February.

__________. 1971. "An Extended Traffic Assignment Model with Application to Two-Way Traffic." Transportation Science. vol. 5, No. 4. November.

__________. 1969. "Network Models for Traffic Assignment." Cornell University Department of Operations Research, Technical Report No. 90. Ithaca, New York, October

Dafermos, S. C. and F. T. Sparrow. 1969. "The Traffic Assignment Problem for a General Network." Journal of Research National Bureau of Standards-B, vol. 73B, No. 2.

Dial, R. B. 1971. "A Probabilistic Multipath Traffic Assignment Model which Obviates Path Enumeration." Transportation Research, vol. 5.

Domencick, T. A., et. al. 1968. "Estimation of Urban Passenger Travel Behavior: An Economic Demand Model." Highway Research Record 238. Washington, D.C. Highway Research Board.

Dunphy, Robert T. and George V. Wickstrom. 1972. "Accessibility: A Tool for Spatial Analysis." Presented at the 1972 Annual Conference of the Urban and Regional Information Systems Association. San Francisco.

Ellis, Raymond H. 1966. A Behavioral Residential Location Model (1966) MS thesis, Northwestern University, June.

"Functional and Detailed Specifications for a Highway Network Equilibrium Program." Prepared for UMTA Planning Methodology and Technical Support Division by Cambridge Systematics, Inc., Cambridge, Massachusetts (forthcoming).

Hudson, James and Steven Lerman. 1974. "Sensitivity Analysis of Equilibrium Flows in Congested Transportation Networks. Paper presented at O.R.S.A/T.I.M.S. meeting, Boston

Jewell, W. S. 1967. "Models for Traffic Assignment." Transportation Research vol. 1, No. 1, May.

Koppelman, Frank. 1974. "Prediction with Disaggregate Models: The Aggregation Issue." Paper in Highway Research Record.

__________. 1975. Ph.D. dissertation, M.I.T.

Kraft, G. 1963. "Demand for Intercity Passenger Travel" in the Washington-Boston Corridor. Systems Analysis and Research Corporation.

Kraft, G. and M. Wohl. 1967. "New Direction for Passenger Demand Analysis and Forecasting." Transportation Research vol. 1.

Landau, Uzi. 1975. "Strategic Analysis Models in Transportation Systems Planning." Ph.D. dissertation in progress, M.I.T.

Lerman, Steven and Nigel Wilson. "An Analytic Model for Predicting Dial-A-Ride System Performance." TRB Special Report 147. Transportation Research Board. Washington, D.C.

Lisco, Thomas E. 1967. "The Value of Commuters' Travel Time: A Study in Urban Transportation." Ph.D. dissertation, University of Chicago.

Manheim, M. L. 1975. Fundamentals of Transportation Systems Analysis. Preliminary Edition, Third Revision.

__________. 1973. "Practical Implications of Some Fundamental Properties of Travel Demand Models." Highway Research Record 422. Washington, D.C.

__________. "Search and Choice in Transport Systems Planning." Highway Research Record 293, Washington, D.C.

Manheim, M. L. and E. R. Ruiter. 1970. "Dodotrans I: A Decision-Oriented Computer Language for Analysis of Multimode Transportation Systems." Highway Research Record 314.

Martin, B. V. and M. L. Manheim. 1965. "A Research Program for Comparison of Traffic Assignment Techniques." Highway Research Record 88.

Martin, B. V., F. W. Memmott, and A. J. Bone. Principles and Techniques of Predicting Future Demand for Urban Area Transportation. M.I.T. Report No. 3.

McFadden, Daniel. 1974. "The Measurement of Urban Travel Demand." Journal of Public Economics, No. 3.

McGillvray, Robert G. 1972. "Binary Choice of Urban Transportation Mode in San Francisco." Econometrica. vol. 40.

McLynn, J. M. and T. Woronka. 1969. "Passenger Demand and Modal Split Models." Northeast Corridor Transportation Project, Report 230, U.S. Department of Transportation.

Murchland, J. D. 1969. "Road Network Traffic Distribution in Equilibrium." Paper for the Tagungueber "Mathematische Methoden in den Wirtschaftswissenschaften." Mathematisches Forschungsinstitut. Oberwolfach, October .

Nestle and Young. 1974. "Short-term Modifications to the Existing CALTRANS Model System to Accept Supply-Sensitive Assignment Techniques and Policy-Sensitives/Demand Models." Working discussion paper, M.I.T., January.

Putnam, Stephen H. 1973. The Interrelationships of Transportation Development and Land Development. Volumes I and II, Report prepared for the U.S. Department of Transportation, June.

Quandt, R. E. and W. J. Baumol. 1968. "The Demand for Abstract Transport Modes: Theory and Measurement." Journal of Regional Sciences. vol. 6, No. 2.

Quarmby, D. A. 1967. "Choice of Travel Mode for the Journey to Work: Some Findings." Journal of Transport Economics and Policy. vol. 1, No. 3, pp. 1-42

Reichman, S. and P. R. Stopher. 1973. "Disaggregate Stochastic Models of Travel Mode Choice." Highway Research Record 369. Washington, D.C.

Ruiter, E. R. 1973. "Analytical Structures" in Highway Research Board Special Report 143 . Washington, D.C.

__________. 1973. "The Prediction of Network Equilibrium: The State of the Art," in Proceedings of the International Conference on Transportation Research. Held in Bruges, Belgium: Transportation Research Forum, June.

Ruiter, E. R. and J. Murchland. 1974. Report to UMTA on equilibration efficiencies, in progress. Cambridge Systematics, Cambridge, Mass.

Stopher, Peter. 1969. "A Probability Model of Travel Mode Choice for the Work Journey." Highway Research Record No. 283. Washington, D.C.

Stopher and T. Lisco. 1970. "Modelling Travel Demand: A Disaggregate Behavioral Approach, Issues and Applications. Transportation Research Form Proceedings.

Strotz, R. H. "The Empirical Implications of a Utility Tree." Econometrica, vol. 25, No. 2, April.

Systems Analysis and Research Corporation. 1963. "Demand for Intercity Passenger Travel in the Washington-Boston Corridor," Report to U.S. Department of Commerce.

Talvitie, Antti and Tom Leung. 1974. "A Parametric Access Network Model." Paper in draft, Oklahoma.

Turnquist, Mark. 1972. A Cost Model for Incremental Traffic Assignment. M. S. thesis, M.I.T.

UMTA Transportation Planning System Reference Manual. 1973. Urban Mass Transportation Administration.

UMTA Transportation Planning System (UTPS) Reference Manual. 1974. UMTA Planning Methodology and Technical Support Division. Washington, D.C., August.

Wardrop, J. G. 1952. "Some Theoretical Aspects of Road Traffic Research." Proc. Institute Civil Engineers. Part II, Road Paper #36. London, pp. 325-378.

Wilson, Nigel, Wayne Pecknold, and Brian Kullman. 1972. "Service Modification Procedures for Local Bus Operations of the Massachusetts Bay Transportation Authority." The Boston Urban Observatory.

Wohl, M. and B. V. Martin. 1967. Traffic Systems Analysis. New York: McGraw-Hill.

Yamamoto, Warren. 1973. A Comparative Analysis of Transit Alternative for Non-Work Trips. M. S. thesis. M.I.T.

EXAMPLES OF COMPUTER APPLICATIONS IN THE TRANSPORTATION FIELD

Faye C. Johnson *

The transportation sector of our economy has integrated the computer into their operations so profoundly that today we find this sector to be one of the most sophisticated and advanced users of computers in the United States. For years American businesses have used the computer in their accounting and administrative functions but only in the last two decades have they put the computer to use in the very heart of their business, i.e., operations. We find in airlines that they now keep all of their passenger reservations on a computer with remote terminals at airports and ticket agencies throughout the United States. They also make extensive use of computers in their maintenance, crew scheduling, and airline ticketing. Railroads, we will see, control their entire car fleets with the use of computers as well as performing very sophisticated analytical simulations to run their operations more efficiently. In motor freight we find their entire rolling stock being tracked by computers as well as their on-line billing and rating.

This paper will address some of the more advanced computer applications in only two modes of our transportation sector and, more specifically, to the hauling of freight in these two modes. These modes are railroads and motor freight companies. The interdependence of our nation's economy on the transportation sector is apparent to everyone associated with transportation. Overall, transportation expenditures, including material inputs to the production of transportation and services, represent about 18%

* Transportation Industry Director, International Business Machines, Corp.

of the United States gross national product and we find these two sectors hauling over 60% of the total freight ton miles by our intercity freight carriers. These figures include both for hire or common carriers and private traffic. Even more significant, we find that railroads and motor freight generate over 94% of all the gross operating revenues from the transportation of goods among the regulated freight carriers.

The fastest growing of these two segments is the motor freight common carrier. With the invention of the internal combustion engine at the dawn of this century, the truck transportation industry has grown in the last thirty years from a very minor entity, in terms of revenue, to the largest single sector of all regulated freight carriers. Several factors can be cited as having contributed to the rapid growth of the trucking industry. These are:

1. The suburbanization of many of the nation's population and employment centers
2. Construction of the interstate highway system
3. The inherent flexibility and convenience of trucks as opposed to other transportation modes

When one looks at the application of computers in these two industries one needs only to look at how these companies spend their available money in the operations of their businesses. Listed below are the percent of expenditures on operating expense of all the motor freight companies in the United States as of 1973.

- o Transportation--46%
- o Terminal operations-22%
- o Equipment maintenance--9%
- o Traffic--3%
- o Other expenses--10%
- o Insurance and safety--4%
- o Administration and general--6%

In railroads as of the end of 1973 we find the following breakdown:

- o Transportation--50%
- o Maintenance of equipment--21%
- o Maintenance of way--18%
- o General expense--7%
- o Traffic--3%
- o Miscellaneous--1%

There is no wonder that both the railroads and motor freight use the computer more extensively to control their transportation expenses than in any other function of their business.

The major difference between the two charts is that while railroads include their yard operations as a part of their total transportation expense, motor freight separates their terminal expense from their transportation expense.

I would now like to address the use of computers in some leading edge applications for the railroad industry.

As mentioned previously, the railroad industry has attained in the last twenty years the status of being one of the most sophisticated users of complex data processing equipment for operational

real-time applications. This position has evolved over several decades whereby they, like many other industries, mechanized many of their batch-related applications through the use of data processing. Examples of these batch applications presently installed by most of the major U.S. railroads are:

- o Waybilling
- o Revenue accounting
- o Centralized payroll
- o Car accounting
- o Disbursements accounting
- o Freight claims
- o Material management
- o Stockholders records
- o Reports to shippers on car status and location

These early batch applications, at the time of their development and implementation, were pressing the state of the art in computer technology and capabilities. A gating factor was computer capability and not railroad requirements. Early data processing, utilizing unit record equipment, had no teleprocessing capabilities and some of the very early computers also lacked this functional capability. When teleprocessing equipment became available, the railroads were among its very first users exchanging information among railroad yards, agencies, and their home office. It wasn't until the third generation of computers became available in the 1960s that the railroads were able to apply these management tools to the very heart of their business, i.e., operations.

We have seen the railroads increase their operating efficiencies over the last thirty years by many means. Some of the more important are:

- o Dieselization
- o Installation of automatic classification yards
- o Installation of microwave communication systems
- o Extensive applications of computers

It is considered by many that the latter may have more impact on greater efficiencies in operations than any of the other developments in the last three decades. It is apparent, with over 1,700,000 freight cars in the U.S., travelling over many thousands of miles of track, the utilization of this valuable asset becomes a major challenge of railroad management.

Today we find that a loaded car is moving only 11% of the time towards its destination. The rest of the time it is empty or sitting in a railroad yard, a repair shop, or at a customer siding being loaded or unloaded. Obviously, improving this 11% payload utilization can mean significant returns to the industry. To be more specific, assuming an average railroad car costs $25,000 each and with 1,700,000 cars in our national inventory, for every 1% increase in the utilization, this industry can have capital expenditure avoidance of $425,000,000. Another way of expressing it--it is equivalent to adding 17,000 additional freight cars to its inventory.

Improving this utilization is the central objective of the use of computers in the railroad industry today. These computer mechanization activities include such applications as:

1. Car Movement Reporting--Keeping the exact status and location of all railroad cars and power on its system and reporting any change of this status to the appropriate points within the railroad that has need of this information
2. Scheduling of Cars--The allocation of available empty cars to shippers, the disposition of empties, and the determination of an appropriate sequence of trains for transporting each loaded and unloaded car from origin to destination
3. Scheduling of Locomotives--The allocation of power in accordance with train schedules, tonnage, and maintenance requirements
4. Scheduling of Trains--Determination of departure and arrival times for each scheduled train; determination of when to operate "extra" trains and to cancel scheduled trains
5. Emergency Responder--Response to systems outages, e.g., derailments, locomotive breakdowns, electrical power and signal failures, bridge washouts, etc.
6. Scheduling of Track (Dispatching)--Control of train movement on all main and branch lines
7. Scheduling of Roadway Maintenance--Allocation of personnel and equipment for inspection, maintenance, and repair of system track and related facilities
8. Scheduling of Car Maintenance--Determination of time and place for inspection, maintenance and repair of freight cars
9. Scheduling of Crews--Selection of trained operating crews in accordance with work contracts, availability, train schedules, and assignment

While this is a very ambitious application list for individual railroads, we find that many of them today are well on their way in their development and implementation. To begin this scenario of applications, a railroad must start with the very complex real-time application--car movement reporting. The objectives of this application are:

1. Reduce clerical burden in yards, agencies and, in some cases, the general office
2. Provide management at all operating levels with complete, accurate, and timely information for improved control and utilization of cars, locomotives, and terminal facilities
3. Establish an historic data base for the eight future real-time applications noted previously

The first phase of car movement reporting can be categorized as the data collection phase. Through the use of lease lines or microwave with data processing terminals installed at several hundred strategic locations throughout the railroad network, many events are reported to a central processing unit as they occur. Some examples of these events are:

- Cars interchanged
- Cars to and from industry tracks
- Cars assigned to arriving and departing trains
- Cars loaded and emptied
- Cars bad-ordered
- Locomotive assignments and status
- Crew assignments

After a network is established to gather this data, then reports are transmitted to appropriate terminals. Most of these reports serve the operating department. Reports generally give status of a car, a group of cars, a train, a yard, etc., or performance of a yard, group of trains, etc. Examples of these reports are as follows:

- Detailed and blocked train consist
- Cars which have high, wide, or heavy loads
- Car inquiries for location and status for manifest information and for advanced car tracing
- Speed-restricted cars
- Car listing by location
- List of trains by route
- Power and caboose reports--on hand and arriving in yards
- Locomotive maintenance due
- Cars delayed in a yard
- Terminal performance reports
- Yesterday's train performance
- Current situation of yard's train and power

After these types of reports are generated, management can issue the following instructions.

Phase I

- Train schedules
- Policy governing blocking rules
- Movement of power from one location to another
- Instructions to hold a car or divert a car
- Special and standing instructions
- Service requests [illegible] forecasts, etc.

After the data collection phase is established, other phases can take place such as:

Phase II

- Traffic functions
- Reports for shippers
- Advance car tracing

Phase III

- Pool inquiries
- Monitor train blocking and train reports

Phase IV

- Dispatcher inputs
- Initial car distribution functions
- Yard summary reports

Phase V

- Revenue data input
- Additional train reports

Phase VI

- Power, caboose, and crew functions
- Car orders and complete car distribution functions and reports

Phase VII

- Holds and diversions

Phase VIII

- Other off-line applications
- Car, locomotive, and caboose statistics
- Demurrage accounting

The car reporting application establishes the data base which is necessary for the railroads to commence the development and implementation of some very exciting additional applications which will further increase the efficiency of their rolling stock. I am categorizing these additional applications under the generic term: railroad operations control. There is widespread agreement that there are two solutions associated with railway freight shipments that can go a long way towards increasing the efficiency of their operations.

1. The ability for the railroads to provide shippers with empty cars for loading of the proper type and grade when they are required and in sufficient quantity.
2. The establishment of dependable and consistent shipper to consignee transit time.

It is clear that the optimum utilization of freight cars requires the solution of extremely complex mathematical scheduling problems. It is equally clear that concepts of mathematical optimization need to be adopted in a systematic manner by the railroad industry in its attempts to solve the varied and interrelated functions of freight operations.

Until quite recently the technology required to formulate, solve, and implement the solutions to the highly complex problems associated with freight operations has not been available.

The necessary ingredients of this technology include telecommunications, on-line high speed data processors having suitable large memory capacity, remote terminals for entry and receipt of control information, as well as efficient algorithms for solving mathematical models in a quasi real-time environment. Much of this technology is

now available. Accordingly, railroads are now taking a fresh look at the railroad operation control system in an effort to achieve maximum benefit for both the railroad industry and its users.

I previously mentioned eight scheduling functions associated with railway operation control and, to repeat, these are:

- Scheduling of cars
- Scheduling of yard operations
- Scheduling of locomotives
- Scheduling of trains
- Scheduling of track
- Scheduling of roadway maintenance
- Scheduling of car maintenance
- Scheduling of crews

While all aspects of freight operations contribute to the quality of service railroads may provide, significant relief of the primary chronic problems may be expected to result in the application of modern technology initially to the scheduling function associated with cars and yard operations. Car scheduling, of course, embraces the movement of both loaded and empty cars. Clearly, these two subproblems are highly interrelated in that typically train traffic is a combination of loads and empties. To facilitate understanding of the special problems of handling of empty cars, it will be helpful to describe briefly the essential scheme for moving cars across a railroad way network.

Freight cars must be moved from their known origin to a prescribed or computed destination (it may be assumed herein that origin and destination are contained within the same railway network).

To accommodate the multiplicity of diverse origin/destination pairs, an operating plan has evolved over the years based upon a hierarchy of trains and strategically situated classifying yards. Cars are switched in the classifying yards into blocks having a common intermediate or final destination in accordance with a railroad blocking policy. These blocks are placed on scheduled trains and transported toward their ultimate destination. The classification policy of a railroad specifies which blocks of cars may be carried on each scheduled train, the order of the blocks, and the handling of the blocks en route. Frequently, the railroad is forced to deviate from its operating plan because of surges in traffic, mechanical failure of equipment, or temporary loss of supporting facilities, i.e., main line, bridges, electrical power, etc. Normal operation is further confounded by system limitations on basic resources such as power, train length and tonnage, crew availability and work rules, yard switching capacity and mainline availability. Deviations from normal operation often trigger cascading effects of significant magnitude.

The input to the car scheduling subsystem would be empty and loaded cars released by users within a given railway network or received at interchange. Output from the system would be car move orders which, in essence, represent trip plans for the movement of each car through the railway network. This system, as I mentioned previously, is supported by their comprehensive car movement reporting and monitoring systems which created the required data base for car scheduling.

The car scheduling subsystem would contain three major functions.

1. Empty Car Disposition--Determines whether an on-line empty car may be considered a candidate for loading. In many instances, regulations prohibit or limit loading of empties and, in effect, prescribe the destination to which a restricted empty must be moved.
2. Empty Car Allocation--Determine the optimum destination for each loadable empty according to the overall system supply, demand, and selected decision criteria.
3. Traffic Scheduler--Determines a suitable trip plan for all necessary car movements in generating appropriate car move orders.

Certain of the above types of problems cannot be resolved within an empty car allocation model or subsystem. However, these problems must be addressed within an overall freight car management system in order to provide necessary support for empty car allocation.

At present, railroads allocate empty cars in a manner similar to the following:

A shipper phones his order to a freight agent. If the order can be filled locally with unassigned empties, this is done. Otherwise, the order is transmitted to a division car distributor. These individuals maintain, to a limited extent, a record of the required number and type of empties needed for their division, along with anticipated releases. Accordingly, the division car distributor knows generally what empties are available for allocation in the shipper's vicinity. He

calls an appropriate yard and places a suitable order for the necessary empties. If a division car distributor is unable to fill an order, he passes it on to the office of the system manager of car utilization. Certain individuals in this office are responsible for allocating specific types of cars, e.g., open top cars (plain flats, coal hoppers, gondolas and chip hoppers). Each week, the system car distributor's office issues a document which contains instructions for disposition of all system and foreign equipment. Daily modifications to this document may be transmitted by telephone.

In a similar fashion, each yard creates a daily disposition sheet which is distributed to car agents, car inspectors, and yard masters.

Generally, empties have a lower priority than loads in that they will be removed first from a train that is too long or heavy. This practice contributes to local yard congestion and may cause shortages at other yards.

In order to provide an effective way to utilize a computer-based technology for the assignment and control of empty railroad freight cars, it is necessary to model the problem analytically. This can be done by applying the resource allocation methodology of operations research.

Factors to be Considered

There are a number of factors which should be considered in the design and development of an optimum-seeking allocation model. These include criteria relating to the accuracy and relevance of the model

within its functional context as well as within its operational context.

Various railroad policies should be accommodated. Among these are:

. Shortage sharing

. Car service rules

. Regulatory constraints

. Per diem minimization

A given customer must ultimately receive a specific car. While one can devise models to handle the assignment of individual cars, it is unclear that this is the most effective approach. It may be more appropriate to group cars into specific "classes" based on car type, condition, special equipment, etc., and assign them to yards based on this classification. The specific assignment of individual cars to nearby customers may be resolved at the yard level, either automatically (by use of a more detailed optimization model), or manually (perhaps with the aid of visual displays), or a combination of both.

Overall control of cars would still be maintained on a systemwide basis by use of the allocation model on a "global" scale, with the specific cars assigned to customers selected from those cars globally allocated to a yard in the customer's vicinity. It would appear that such a two-level approach (aggregating at the higher level and "exploding" the aggregated solution at the lower level) is more reasonable.

This last point relates as well to the operational use of the model. To be effective, the model must be used regularly on a semi-real-time basis. [illegible] the problem must be solved depends, in part,

upon the dynamic character of the problem data. This may be several times a day or as infrequently as once a day.

While empty cars must be allocated throughout a railway system, there is a natural division of this problem into two parts. One deals with the global aspects of empty car allocation--the systemwide distribution of cars between major yards. The other part of the problem deals with the local aspects, i.e., ensuring that individual shipper's requests for empties are satisfied, as well as determining where released empties must be sent.

In terms of actual railroad practice, the differentiation between these two problems (global and local) is important. Consideration of the global problem leads to simplifications since one need not consider all the detailed differences between car types and can aggregate similar types for routing between major yards. Nevertheless, at the local level it may become necessary to distinguish between cars.

However, local problems are separable, i.e., are independent of each other; thus, the local car distributor's knowledge can be utilized to facilitate solution of the problem. It will often be possible to take advantage of the distributor's familiarity with the shippers in his area regarding such questions as the condition of a released car, or whether substitution of one type of car for another is acceptable.

To formalize the distinction between the global and the local problem, the concept of a car service area will be introduced. A car service area is that "territory" serviced by a single major yard; while there may be secondary supply points within the car service

area from which cars may be supplied to customers in the area, any car which comes from outside the car service area must enter the area through the major yard.

Major yards are linked by through trains, while individual shippers are served by local trains or yard engines which typically run between two supply points, or out and back from a single supply point.

Car service areas may be considered as a mechanism whereby local information can be aggregated for systemwide use. That is, the net supply or demand in the service area can be determined and associated with that area's major yard. Thus, in terms of the global model, systemwide empty car allocation may be based on the net supply or demand at each of the major yards. Having solved the allocation problem at the global level, solution of the associated local problems can be obtained by "exploding" the global solutions. In the event that there are alternative ways of supplying or receiving cars within a car service area (for example, when there are secondary supply points), an additional optimization will be required. Because of the way in which car service areas are defined, treating the problem in this two-level fashion (global-local) and optimizing individually, one can achieve the same overall optimal solution as if all the problem aspects (both global and local) were combined into one large problem.

In many cases, the local problem need not be solved by an optimization technique for there may be no alternatives to choose between. For example, this occurs when a customer on one branch in the car service area must receive his cars from a single supply point and, when cars are released, must return to the supply point. In these

instances, it is sufficient to keep a detailed list at the supply point of requirements and releases; these cars will be moved on the first available train consistent with the release/requirement time. If there are insufficient cars available to satisfy requirements within a car service area, the global problem solution will provide for empties to move to the appropriate major yard. Similarly, if surplus cars are available, the global solution may specify they be sent to other major yards.

The Allocation Model

The allocation model utilizes a two-level scheme, with one level treating the global or systemwide aspects of the problem, and the other dealing with its local aspects. For the global problem, cars in a class would be aggregated into groups by type, location, and time of availability such that it would not be necessary to distinguish between cars within a group. For the local problems, specific cars in a class may be considered directly in the model. Thus, a specific car, if not needed locally, would be grouped and allocated in a global problem, and then assigned specifically in a subsequent local problem run.

The determination of the final destinations for empty cars is possible through the use of simulation techniques. Dynamic scheduling of cars and trains would be the end result whereby railroads could meet their overall objectives of railroad operating control systems.

The most important aspect of a viable operating control system is to provide better controls over the railroads' assets. It will provide:

1. Better control of trains because advanced information would be available on the exact volume of traffic to be moved
2. Better control of empty and loaded cars--a determination will be made as to which cars will be moved in what sequence along what route and on what train
3. Better control of car distribution--disposition instructions are part of the computer programs and would direct the movement of empty cars. The computer would signal any failure of an empty to move as planned

The development of an intraline operating system is required before a nationwide interline system can be developed. Railroads are moving soundly towards their achievement of the first phase. The Association of American Railroads with their TRAIN II system is providing the basic framework for a national car interline control system.

When this event occurs, it is the writer's feeling that the occurrence will be recognized as the single most significant happening in the railroad industry during its entire history.

<u>Motor Freight Industry</u>

Now, let's turn our attention to the motor freight industry. This sector is the lifeline of goods for our country. It's a business that never closes its doors without affecting every industry and every consumer--a business that's built on giving a mobile society the freedom it needs to prosper.

The ever-rising cost of labor and equipment, competition, and unexpected events like our recent energy crisis, present constant and changing challenges.

For carriers, external, uncontrollable factors such as fuel prices, labor contracts, financing, and interest rates create problems in many of their operational areas. But carriers do prosper in spite of these external factors when management improves its handling of its own internal operations.

Probably the most important areas that carrier management is addressing are its freight terminal and linehaul operations. By improving its impact on these areas, management is accomplishing three basic objectives: service, productivity and control.

Service is the carrier's competitive edge. Providing better shipment status information and insuring accurate billing and invoicing on a timely, consistent basis are the keys to improved customer satisfaction.

Productivity for motor carriers is coming from the increased productivity of people and equipment. The areas which are being addressed include linehaul equipment availability and utilization, dock worker productivity, reduction of empty movements, and improved load factors.

Control of operations means control of equipment in the areas of balance, interchange, maintenance, and licensing. It means control of revenue by eliminating lost bills and transcription errors, and providing positive delivery verification. It means control of freight to reduce over, short, and damage; reduce inactive freight; and facilitate tracing.

In view of the demands placed on management today, new tools and techniques have been developed to sustain the profit and growth rates of the 1960s. One of the tools that many carriers have looked to is the computer.

It has been used in the motor freight industry for some time. Today it is possible to interface a computer to a communications network and capture information as events occur. As a result, information is being retrieved on demand, information that's timely and accurate, that allows decisions on service and productivity to be made on a real-time basis.

Implementation of computer-based teleprocessing systems has typically been an expensive and time-consuming proposition when undertaken by an individual carrier. In working with the motor freight industry, IBM has developed a system of application programs to answer the operational needs of service and productivity, while avoiding the duplication of substantial dollar expenditures and effort.

Other similar systems are available; however, because of my familiarity with the system we developed, I will confine my remarks to our system as illustrative of what carriers are doing in these advanced application areas.

We call our system FERST/VS--Freight and Equipment Reporting System for Transportation/Virtual Storage. FERST/VS is a teleprocessing system designed to provide the management of a motor freight company with timely and accurate information on the movement of freight and linehaul equipment throughout their system.

FERST/VS offers four application programs to assist operations management: message switching, equipment control (including a shipment tracing function), freight billing, and rating rate audit. These programs help improve customer service, increase the productivity of equipment and personnel, and give motor carriers better control of operations.

Let's now discuss each of these areas in detail. Morning reports, terminal performance and statistical reports, hot shipment advice, or over, short, and damage reports are entered and transmitted effectively between terminal locations and the central office. This is called the message switching application.

The motor freight industry has recognized that the key to success lies in the control of terminal and linehaul operations. Efficient scheduling of linehaul and city equipment and planning of terminal resources are critical if the use of manpower and equipment resources is to be maximized.

If the dispatcher wants a list of all tractors or trailers at a given location, he can enter an equipment inventory inquiry. He is presented with a list and status of that equipment.

Changes to equipment status are reflected in the report generated from the equipment inventory request. This information is as current as the entries made through the communications devices.

One of the high points of equipment control is the information it supplies to terminal management on schedules which are currently en route to their location. Terminal personnel retrieve information on schedules destined for them within a specified number of hours (e.g., next sixteen hours) or on total schedules moving in the system with freight destined for their terminal.

The report generated as a result of this inquiry gives the terminal personnel estimated time of arrival by schedule, weight/cube/number of bills, and any specific load information entered at dispatch, trailer close, or billing time.

The schedules en route report has many potential uses including:

- Preplanning dock and city equipment requirements
- Preplanning trailer spotting at the dock based on large mark information
- Anticipating equipment availability at a given location
- Providing accurate information on inbound tonnage and shipments
- Projecting dock crew workloads

Once a shipment has been recorded in the system's data base, the current status and location of that shipment is retrieved by entering an inquiry. The operator receives an immediate response from the computer showing the location and status of the shipment.

Revenue, shipments, and equipment must all be considered in the control system for the operation of a motor carrier. Efficient utilization of equipment control system which is integrated into the freight billing system.

Let's now look at the freight billing application and see how it interfaces to the equipment control facility in order to provide a full-function motor freight operations control system.

Equipment control is a computerized system designed to assist dispatching, keep real-time equipment inventories, and provide shipment tracing and advance schedule and load information.

Equipment control application functions are broken into the following three areas:

1. Equipment movement reporting, which consists of simple input messages used to report changes in status and location of linehaul equipment.

2. Inquiry and tracing, which allows the carrier to inquire as to the current location of equipment, equipment inventories at any terminal location, schedules en route between locations, and the last reported status of a particular shipment.
3. Shipment control which gives the carrier the ability to report changes in status and final disposition of individual shipments.

Now, let's take a more detailed look at some of the features of equipment movement reporting and how they are helping the motor carrier's operation. For example ...

When a schedule (that is, tractor, trailer combination) is ready to move, a dispatch message is entered. The message reports the movement of that schedule to the computer and updates the equipment files to show those pieces of equipment in transit between terminals.

In order to report trailer interchange between carriers, the program provides a trailer rental and interchange message. Once the information regarding the interchange is entered into the computer, the carrier keeps track of the time when the trailer was taken on-line or put off-line and the return point. This information facilitates tracking interchange activity and calculating accurate per diem charges.

The equipment movement reporting facility also allows the carrier to report on changes in equipment status. Examples would be a tractor placed in the shop for maintenance or a trailer spotted at a customer location for loading or unloading.

Let's now look at the inquiry and tracing facility. Here are some of the facilities it provides that are helping carriers improve their linehaul and terminal operations.

Freight billing is divided into these functional areas:

- Bill entry
- Manifesting
- Delivery reporting
- Management reports

Many edits are performed on the bill before it enters a system. For instance, pro numbers are checked for validity, destination codes are checked to be sure they are valid, extensions and totals are checked for accuracy. Conversational editing of the bills at this stage helps insure accuracy because the person creating the bill can make corrections while the original bill of lading is at hand. This type of bill entry eliminates the need to keypunch bills for entry into a revenue account system, avoiding a step where transcription errors can occur.

Since all of the information from the freight bill is available from the data base, it is possible to do manifesting and to produce more detailed reports to be used as advanced planning by terminal management.

The two primary documents produced by the manifesting section of the freight billing are the tracing manifests and the trailer summary manifests.

Tracing manifests are available to terminal locations on request for both inbound and outbound shipments. Inbound manifests are sorted into alphabetical sequence by consignee; outbound

manifests are sorted into alphabetical sequence by shipper. Terminal personnel uses the tracing manifests in conjunction with the trace inquiry to answer all shipment inquiries submitted by either shipper or consignee on a particular pro number.

The trailer summary manifest is an extension of the schedules en route report which is produced in the equipment control program. As a result of the freight billing operation, weight, destination, and additional comments are available. The trailer summary manifest is of particular significance to break bulk operations because it summarizes shipments, pieces, and weight by ultimate destination. Terminal management uses the trailer summary manifest in planning its city pickup and delivery requirements.

Prior to the arrival of freight at destination, full freight bill delivery sets are printed on request at the terminal. The delivery sets are used for the delivery of the freight. They may also be the documents used to control stripping of the linehaul trailer and pickup and delivery loading.

The freight billing system allows the carrier to report the final disposition of a shipment via a delivery message. If the carrier chooses, this message will trigger automatic invoicing. It may also be used to record over and short information.

Once a system has been implemented to do overhead freight bill transmission, the centralization of the rating and rate audit functions is a logical next step.

The rating rate audit application package provides the three functions required of such a system:

1. Rating--separated from bill creation and supported at one or more central facilities.
2. Rate audit--where the billing system selects those bills to be audited automatically based on user-defined parameters.
3. Rate quote--which allows terminal personnel at remote locations to submit requests for rate quotations through their communications devices.

With rating rate audit installed, the billing clerks enter skeleton bills at the origin terminal. Rating personnel at a central facility are presented sequentially with the unrated bills grouped by tariff bureaus. The rate clerk then enters the appropriate rate. The rating program will then perform the extensions and total the charges.

The rate audit automatically selects those bills which are to be audited. The invoicing of bills queued for auditing is prevented until the audit procedure is completed. Many carriers select the weight and charge parameters that determine which bills are selected for auditing.

The rate quote facility allows terminal personnel to enter requests for rate quotations on their communications devices. The requests are presented to rating personnel at the central facility. They enter the appropriate response, and the request is sent back to the origin t

Among the potential benefits to be realized through the use of the rating/rate audit programs are:

- Reducing the wait time for shipments
- Alleviating pressure on rate clerks--reducing their errors
- Allowing a greater degree of rate specialization through centralization of personnel
- Reducing the number of rate clerks

In addition to the on-line capabilities we have discussed, these programs generate a number of summary and statistical reports that help a carrier measure performance and more effectively control his operation.

Inactive freight reports help terminal management control over, short, and damage, and improve customer service.

To summarize the potential benefits of these on-line applications, let's recap the areas of service, productivity, and control.

Under service, the shipment tracing inquiries and reports help to answer customer requests more quickly. The equipment control portion helps to balance equipment, keep track of special equipment, and make it available where it is needed, when it is needed. Freight bill transmission helps move the freight faster, reduces over-and-without bill situations and the on-line editing of bills reduces biller errors. Delivery times recorded provide a record for management of how schedules are being met. The inactive freight reporting helps terminal management control over and shorts.

Under productivity terminal managers can anticipate manpower and city equipment needs with the schedules enroute and trailer summary reports. Central dispatch and terminal managers can improve load factors, increase utilization, and reduce inactive equipment with the

accurate, up-to-date equipment inventory. Biller production can be improved. Rating/rate audit further improves biller production, because billers no longer have to enter rates and charges on each bill. Centralized rating also helps clerks become more productive.

In the area of control these applications offer better control by providing a data base of information on the carrier's operations when and where needed. That can mean better decisions made, potential problems anticipated, and action taken before potential problems become real ones.

Information is provided to:

- Terminal management with--
  - Schedules en route reports
  - Empty movement reports
  - Inactive freight reports
  - Biller production reports
- To linehaul operations with--
  - Equipment inventories
  - Tractor cycle and relay reports
  - Inactive equipment reports
- To accounting with--
  - Master bill file verify and balance reports
  - Delivery verification
- To data processing management with--
  - Terminal usage reports
  - Communication circuit analysis
  - System error incidences

And it also provides information to executive management who can review the same information to measure relative terminal performance, fleet utilization, biller and dock productivity, and many other facets of his business.

I hope the preceding computer application examples have given you some feeling for the considerable extent to which our transportation industry has applied the computer in their day-to-day operations.

While most of the uses have been in the intramodel area only, it will not be too many years away before intermodel exchange of information and control becomes commonplace. Perhaps some time in the future we can have a worldwide transportation control system.

COMPUTER APPLICATION IN THE ALLOCATION OF AIRLINE RESOURCES

Morton Ehrlich *

Airlines today, at least in the United States, operate in an environment characterized by:

- Limited resources
- Rising costs
- Restrictive labor contracts
- Government regulatory policies
- Increased competition
- Economic sluggishness, depressing traffic growth

Within this environment, the challenge is not necessarily one of viability, but one of survival. Classical solution systems are no longer applicable to today's problems. New and advanced methodologies and techniques must be developed to respond to current competitive and market pressures.

This paper addresses the major problem area of the practical allocation of aircraft resources as a function of market demand and fleet mix, and discusses the development of new systems designed to respond to the problem.

There are three key phases which dictate, to a large extent, the allocation of airline resources:

- market planning, which defines market strategies and revenue generation
- schedule planning, which converts market strategies to specific flights, and coordinates implementation
- flight crew pairing and allocation, which attempts to optimize the use of the human resource

* Vice President of Planning, Eastern Airlines, Inc.

All these phases converge into schedule development, the master control of the consolidated allocation process.

Within each phase, computer systems provide support:

- A marketing information system using data base management software supports the market planning function. The data base contains several years of proprietary eastern, competitive, and industry traffic and operational statistics.
- A frequency planning model is used to conduct preliminary evaluations of future schedules. Results recommend the optimal sizing and frequency of utilization of aircraft, by fleet type.
- An online flight crew pairing system provides crew management with the capability to interact with computer-stored files, and update the crew schedule with evolving changes to the general flight schedule.
- An online schedules development system provides the capability for online updates to the master flight schedule. This system provides overnight turnaround of schedule listings reflecting the latest resource allocation judgements.

Before discussing the resource allocation function, let me give you an overview of the facilities and hardware used to perform the allocation process.

EASTERN'S COMPUTER COMPLEXES

Eastern has three major computer centers:

- At Charlotte, North Carolina, the approximate geographical center of the domestic route structure, are three UNIVAC 494 computers with supportive peripheral hardware and software. This center monitors and supports the day-to-day operation of the airline, through an extensive, high-speed communication network operating in a real time, online environment. Statistics collected throughout the day are transmitted between 2:00 A.M. and 4:00 A.M. over a high speed data link to the administrative center in Miami for report production within a batch environment.

- In Miami, there are two other major complexes: one, supporting the systemwide reservations network; the other, providing requisite processing for the administrative, financial, and engineering and maintenance requirements.

Two IBM 370-195s and two IBM 360-65s configure the reservations complex. In addition, they

- Support the automated passenger processing systems at our major airports and ticket offices. This system makes available to the counter agent within a real time framework:
  - . Reservations status for all flights
  - . The capability to offer seat selection within an automated environment
  - . The capability to generate automatically tickets from the reservations record
  - . Flight operational statistics
  - . Provides total processing support for several regional airlines' reservations requirements

The system also communicates with the administrative computer complex via a high-speed data link.

The administrative complex is comprised of four IBM 360-65s, one IBM 360-30, and several smaller computers used either as front ends for high-speed communication networks or support for interactive systems oriented to a single function (Exhibit I). Systems discussed in support of the resource allocation process all reside within the administrative complex. Now let me start developing the application framework.

MARKET PLANNING

The market planning function has the primary responsibility for revenue generation. Activity groups comprising the function are responsible for developing market strategies, monitoring and analyzing market traffic performance, and coordinating product-related input to Sales and Advertising programs.

Market Strategy Development

On a monthly basis, market strategy development monitors (Exhibit II):

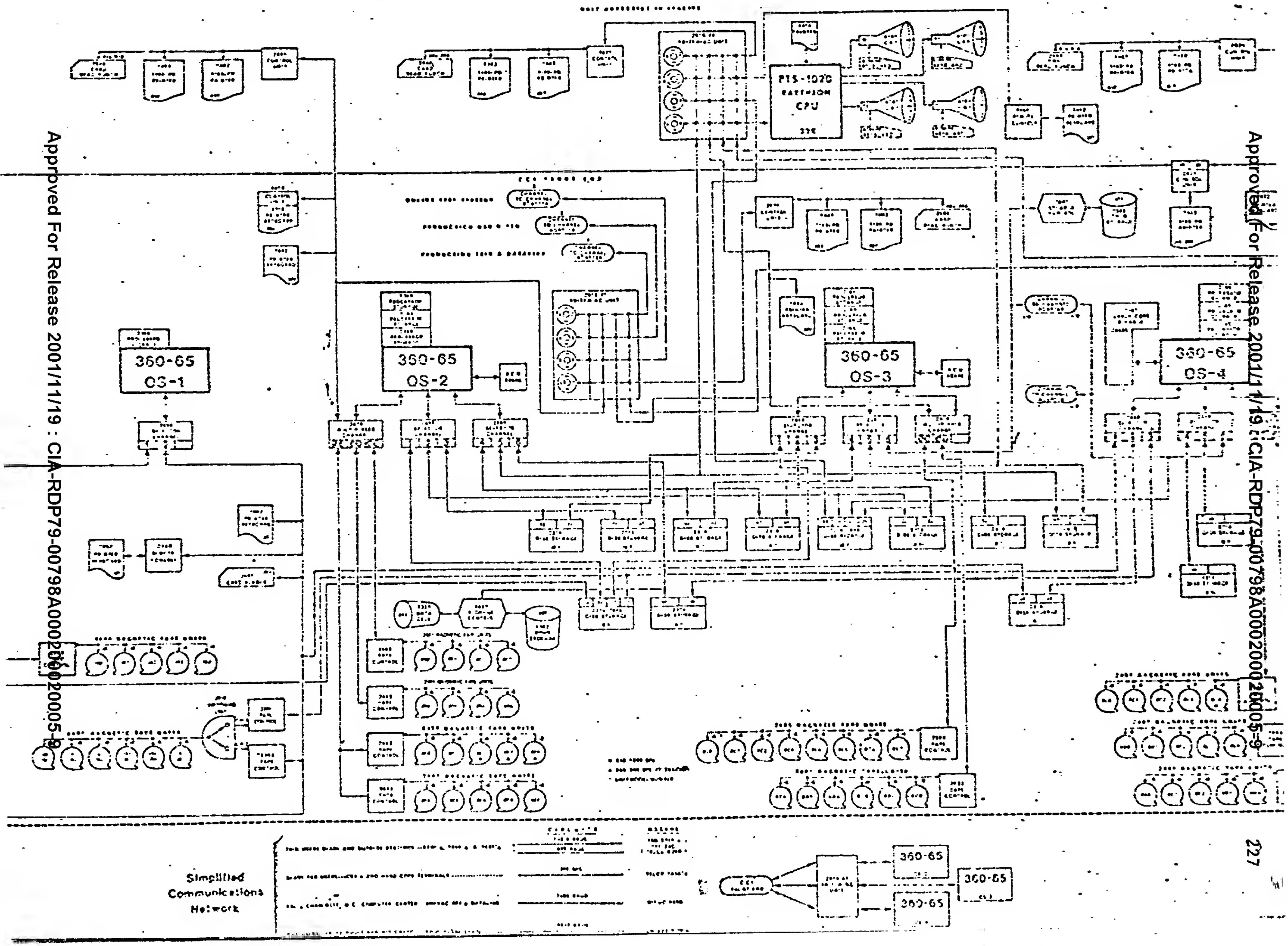
- Industry traffic and operational statistics
- Competitive actions
- Eastern's performance for each market on the eastern system

From three to six months prior to the schedule-planning process, the top 400 markets are subjected to analysis. At this point in the process a demand forecast of industry traffic is made for each market. The results are then used in a forecast algorithm to allocate industry

EXHIBIT I

COMPUTER SCIENCES - COMMERCIAL COMPUTER SYSTEMS

EQUIPMENT SCHEMATIC



Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9

*** DATA BOOK ***

1974 OPERATING PLAN
BASED ON- OPERATING PROPOSAL DATED AUG. 21
ATLTLH

	PERCENT INDUSTRY OED	INDY OED	ONED INDY	MARKET SHARE	SEAT SHARE	PERCENT EASTERN OED	EALPS CCC	YXC CCC	NXT CCP	HRO SGP SGP	NED SGR	PERCENT ONBOARD	SK LOAD FACTOR	EA-SK SEATS	COMP SEATS	PERCENT TOTAL SEATS	EA-SK A-STP	COMP A-STP	1ST PAGBOARDING	PERCENT 1ST CITY BOARDING	2ND CITY BOARDING	PERCENT 2ND CITY BOARDING
72AAA	+20.50	116	0	+56.74	+55.55	+43.58	112	0	0106	53	274	+25.86	+66.36	513	0	+22.27	4.81	.00	258	+25.58	225	+29.71
72AAA	+15.39	135	0	+56.74	+55.96	+15.50	130	0	0112	46	289	+5.30	+59.64	484	0	+17.15	5.01	.00	268	+3.51	264	+15.25
74MLP	+5.60	145	0	+57.41	+100.00	+9.35	142	0	0128	41	311	+7.55	+63.72	488	0	+.62	5.00	.00	287	+7.24	291	+9.99
7212A	+21.30	130	0	+56.52	+100.00	+21.15	126	0	0130	46	302	+11.03	+64.11	471	0	+11.61	4.92	.00	280	+7.28	275	+20.78
7301A	+13.51	126	0	+56.82	+100.00	+14.02	122	0	0 75	68	265	+5.58	+52.53	455	0	+21.52	5.07	.00	245	+2.08	244	+17.37
7302A	+17.55	138	0	+56.37	+100.00	+17.70	133	0	0103	35	271	+3.44	+56.10	483	0	+17.52	5.02	.00	256	+3.64	244	+5.51
7303A	+9.45	133	0	+56.40	+100.00	+8.94	134	0	0118	38	290	-.69	+58.94	492	0	+18.84	5.02	.00	260	-9.41	265	+8.91
730-1	+12.18	143	0	+57.20	+100.00	+12.80	139	0	0136	42	317	+7.09	+65.62	483	0	+16.11	5.05	.00	259	+15.44	286	+18.67
7312P	+7.65	140	0	+57.25	+100.00	+8.73	137	0	0140	48	325	+7.62	+66.55	486	0	+3.61	5.00	.00	300	+7.14	307	+10.04
-01P	+11.50	141	0	+57.37	+100.00	+13.11	138	0	0115	36	287	+6.30	+58.91	488	0	-1.42	5.00	.00	265	+8.16	268	+5.54
02P	+7.25	148	0	+57.29	+100.00	+8.27	144	0	0111	37	291	+7.38	+59.63	488	0	+1.04	5.00	.00	274	+7.03	268	+9.84
403P	+7.19	149	0	+57.31	+100.00	+8.21	145	0	0127	40	312	+7.55	+63.93	488	0	-.81	5.00	.00	278	+6.52	256	+10.04
10 P	+9.07	156	0	+56.79	+100.00	+6.63	151	0	0144	44	339	+6.94	+65.46	488	0	+1.04	5.00	.00	320	+7.02	315	+10.14

projections among Eastern and competitors. Several iterations of the statistical forecast are required to test multiple competitive assumptions.

When finalized, each of the markets will have been assigned a priority based upon its revenue-generating potential. The priority is later used to determine alternative tactical plans when recommended priorities cannot be accommodated because of resource or utilization constraints.

Traffic Performance Analysis

In depth analysis of Eastern's top 400 markets include enplanements, onboard traffic, O&D traffic, seats, departures, load factor, seat share, and trip share. Two reports serve as the basis for the analysis:

- A market traffic performance analysis (Exhibit III) reports traffic at the non-directional market level with summaries of sixty economic and market group entities. This report is oriented to sales and advertising functions, and provides the option of adjusting strategies and programs to conform with market performance.
- A directional traffic performance analysis (Exhibit IV) reports the same data elements as the market level report; however, it reflects traffic performance at the directional station pair level, with summaries at the station and regional levels. This report is used during the schedule development cycle to fine-tune strategic and tactical plans scheduled for future implementation.

YR/YR PERFORMANCE ANALYSIS OF ONBOARD SEATS TRIPS AND ONBOARD PSGRS FOR FEBRUARY 01 TO 15 74/75

GROUP 27

EXHIBIT III

JET CHOICE PIEDMONT

SEG	ONBOARD SEATS HIST	ONBOARD SEATS ACTUAL	ONBOARD SEATS VARIANCE ABS	ONBOARD SEATS VARIANCE PERCENT	ONBOARD TRIPS HIST	ONBOARD TRIPS ACTUAL	ONBOARD TRIPS VARIANCE ABS	ONBOARD TRIPS VARIANCE PERCENT	ONBOARD PASSENGERS HIST	ONBOARD PASSENGERS ACTUAL	ONBOARD PASSENGERS VARIANCE ABS	ONBOARD PASSENGERS VARIANCE PERCENT	LOAD FACTORS HIST	LOAD FACTORS ACTUAL	LOAD FACTORS VARIANCE ABS	LOAD FACTORS VARIANCE PERCENT
[illegible]	175	317	142	81.14	2.3	3.0	1.0	50.00	50	65	15	30.00	28.57	20.50	8.07-	28.25-
[illegible]	586	657	71	13.62	5.0	6.0	1.0	20.00	333	307	74-	14.32-	65.36	45.33	19.03-	29.12-
[illegible]	486	626	140	28.81	5.0	5.2	.2	4.00	358	283	75-	20.51-	73.25	45.21	28.04-	38.28-
[illegible]	577	652	75	13.00	5.5	5.5	.0	.00	453	372	81-	17.88-	78.51	57.09	21.45-	27.32-
[illegible]	91	89	2-	2.20-	1.0	1.0	.0	.00	45	41	4-	8.89-	49.45	46.07	3.38-	6.84-
[illegible]	0	93	93	.00	.0	1.0	1.0	.00	0	69	69	.00	.00	74.19	74.19	.00
[illegible]	240	279	39	12.50	3.0	2.8	.2-	6.67-	133	116	22-	15.94-	55.65	41.58	14.07-	25.28-
[illegible]	100	97	3-	3.00-	1.0	1.0	.0	.00	24	21	3-	12.50-	24.00	21.65	2.35-	9.79-
[illegible]	59	59	0	.00	1.0	1.0	.0	.00	45	31	14-	31.11-	50.56	34.83	15.73-	31.11-
[illegible]	88	86	2-	2.27-	1.0	1.0	.0	.00	39	33	6-	15.38-	44.32	38.37	5.95-	13.43-
[illegible]	261	281	20	7.66	2.9	3.0	.1	3.45	99	78	21-	21.21-	37.93	27.76	10.17-	26.81-
[illegible]	192	359	123	57.55	2.4	2.9	.5	20.83	142	146	1	.69	73.96	47.25	26.73-	36.13-
[illegible]	89	131	44	49.44	1.0	1.9	.9	90.00	58	58	0	.00	65.17	43.61	21.56-	33.09-
[illegible]	0	199	199	.00	.0	2.1	2.1	.00	0	70	70	.00	.00	35.18	35.18	.00
[illegible]	264	328	64	24.24	3.0	3.0	.0	.00	141	136	5-	3.55-	53.41	41.46	11.95-	22.37-
[illegible]	0	6	6	.00	.0	.1	.1	.00	0	2	2	.00	.00	33.33	33.33	.00
[illegible]	216	178	38-	17.59-	2.0	2.0	.0	.00	102	56	46-	45.10-	47.22	31.46	15.76-	33.38-
[illegible]	2	0	2-	100.00-	.0	.0	.0	.00	1	0	1-	100.00-	50.00	.00	50.00-	100.00-
NYC	693	684	9-	1.30-	6.5	6.2	.3-	4.62-	415	332	83-	20.00-	59.88	48.54	11.34-	18.94-
PHL	279	295	16	5.73	2.9	2.9	.0	.00	190	146	44-	23.16-	68.10	49.49	18.61-	27.33-
PIT	359	420	61	16.99	4.0	4.0	.0	.00	202	196	6-	2.97-	56.27	46.67	9.60-	17.06-
[illegible]	2	0	2-	100.00-	.0	.0	.0	.00	2	0	2-	100.00-	100.00	.00	100.00-	100.00-
[illegible]	111	96	15-	13.51-	1.0	.9	.1-	10.00-	39	31	8-	20.51-	35.14	32.29	2.85-	8.11-
[illegible]	270	343	73	27.04	3.0	3.5	.5	16.67	209	192	17-	8.13-	77.41	55.98	21.43-	27.68-
NYC	351	237	114-	32.48-	3.5	2.4	1.1-	31.43-	173	128	45-	26.01-	49.29	54.01	4.72	9.58
PIT	126	138	12	9.52	1.1	1.2	.1	9.09	72	61	11-	15.28-	57.14	44.20	12.94-	22.65-
[illegible]	177	223	46	25.99	2.0	2.0	.0	.00	128	96	32-	25.00-	72.32	43.05	29.27-	40.47-
NYC	132	132	0	.00	1.0	1.0	.0	.00	85	59	26-	32.95-	66.67	44.70	21.97-	32.95-
[illegible]	95	88	7-	7.37-	1.0	1.0	.0	.00	54	36	18-	[illegible]	60.67	39.58	21.09-	34.76-
[illegible]	433	437	4	.92	4.5	4.0	.5-	11.11-	258	221	37-	14.34-	59.58	50.57	9.01-	15.12-
RIC	377	473	96	25.46	3.9	4.9	1.0	25.64	241	213	28-	11.62-	63.93	45.03	18.90-	29.56-
[illegible]	111	118	7	6.31	1.0	1.0	.0	.00	85	56	29-	34.12-	76.58	47.46	29.12-	38.03-
[illegible]	435	451	16	3.68	4.0	4.0	.0	.00	286	235	51-	17.83-	65.75	52.11	13.64-	20.75-
ALL	7413	8572	1159	15.63	75.2	81.4	6.2	8.24	4521	3890	631-	13.96-	60.99	45.38	15.61-	25.59-

MARKET PLANNING & FORECASTING
YEAR/YEAR TRAFFIC PERFORMANCE ANALYSIS
APRIL 1975/1974

EXHIBIT IV

REGION OF ORIGIN ORIGIN STATION

CENT-NORTH FLA TPA

DES STA	O&D-PSGR CUR YEAR	O&D-PSGR CHNG ABS	ONBRD-PSGR CUR YEAR	ONBRD-PSGR CHNG ABS	BOARDINGS CUR YEAR	BOARDINGS CHNG ABS	LOAD-FACTOR CUR YEAR	LOAD-FACTOR CHANGE ABS	0-STOP-SEATS CUR YEAR	0-STOP-SEATS CHNGE ABS	0-STOP-TRIPS CUR YEAR	0-STOP-TRIPS CHANGE ABS	1-STOP-TRIPS CUR YEAR	1-STOP-TRIPS CHNGE ABS	COMP-SEATS CUR YEAR	COMP-SEATS CHNGE ABS	COMP-0-ST-TRP CUR YEAR	COMP-0-ST-TRP CHANGE ABS	CUR-YEAR SEAT SHARE	CUR-YEAR TRIP SHARE
[illegible]	10	-2					*	*			.00	.00	.00	.00			.00	.00	*	*
[illegible]	153	-10	515	-22	511	47	66.70	-14.66	772	112	6.20	.13	1.00	-1.00	2047	744	11.00	2.00	26.7	35.2
[illegible]	4						*	*			.00	.00	.00	.00			.00	.00	*	*
[illegible]	24	24					*	*			.00	.00	2.00	2.00			.00	.00	*	*
[illegible]	82	5	31	-50	31	-31	73.80	-17.21	42	-47	.47	-.53	1.50	1.50			.00	.00	100.0	100.0
[illegible]	3	-1					*	*			.00	.00	.00	.00			.00	.00	*	*
[illegible]	27						*	*			.00	.00	1.00	1.00			.00	.00	*	*
[illegible]	93	-44	90	-39	65	-57	60.81	-5.00	148	-48	1.43	-.57	.87	-.13	336	-81	3.00	.00	31.9	32.7
[illegible]	15	5					*	*			.00	.00	.00	.00	99	-25	1.00	.00	.0	.0
[illegible]		-2					*	*			.00	.00	.00	.00			.00	.00	*	*
[illegible]	41	-12	40	-43	49	-43	61.53	-.41	65	-69	.53	-.14	.00	.00	395	-89	2.00	-1.00	14.5	33.3
[illegible]	21	5					*	*			.00	.00	1.00	1.00			.00	.00	*	*
[illegible]	69		110	14	97	1	85.93	-9.11	128	27	1.43	.40	.00	.00			.00	.00	100.0	100.0
[illegible]	1	1	2	2	2	2	33.33	*	6	6	.07	.07	.00	.00	278	278	2.00	2.00	.0	.0
[illegible]	82	82	154	-7	182	-9	73.96	3.47	233	-20	2.00	.00	.00	.00	270		2.00	.00	46.5	50.0
[illegible]	33	-2	71	-3	46	-1	39.68	-16.60	178	47	2.00	1.00	.00	.00	492	-28	4.00	.00	36.5	33.3
[illegible]	155	92	182	92	170	89	68.93	-17.60	264	160	2.53	1.43	.00	.00	393	-85	2.00	-1.00	40.3	56.2
[illegible]	8	-2					*	*			.00	.00	.00	.00			.00	.00	*	*
[illegible]	162	-45	149	-25	149	-25	71.63	2.59	208	-44	1.03	.00	.47	-1.53			.00	.00	100.0	100.0
[illegible]	96	13	152	13	110	23	58.23	-6.42	261	46	2.93	.93	.00	.00	380	-109	1.97	-2.00	40.5	60.1
[illegible]		-3					*	*			.00	.00	.00	.00			.00	.00	*	*
[illegible]	13						*	*			.00	.00	1.00	1.00			.00	.00	*	*
[illegible]	9	-3					*	*			.00	.00	.00	-1.00			.00	.00	*	*
[illegible]	2	2					*	*			.00	.00	.00	.00	135	135	1.00	1.00	.0	.0
[illegible]	52	-6	59	59	54	54	64.93	*	91	91	1.03	1.03	.00	-3.00			.00	.00	100.0	100.0
[illegible]	150	40	214	96	177	70	73.53	3.34	291	110	3.10	1.17	.00	-.50	649	92	5.00	.00	31.6	39.0
[illegible]	2	2					*	*			.00	.00	.00	.00			.00	.00	*	*
[illegible]	71	-23	64	-28	54	-38	74.41	3.65	86	-44	.93	-.10	.00	.00	139	-115	1.00	.00	35.8	50.7
[illegible]	21	7	119	81	51	13	40.20	23.54	296	68	3.00	2.00	.00	.00	929	-465	6.00	-3.00	24.1	33.3
[illegible]	2	2					*	*			.00	.00	.00	.00			.00	.00	*	*
[illegible]	129	16	157	-17	145	-3	49.76	-22.10	404	102	4.30	1.10	1.53	1.53	1650	-157	9.26	-1.00	15.5	31.0
[illegible]	2	-1	2	-1	2	2	66.66	-33.34	3		.03	.00	.00	.00	320	90	1.00	.00	.0	.0
[illegible]	5	5	30	30	13	13	19.35	*	155	155	1.50	1.50	.00	.00	135	135	1.00	1.00	53.9	60.4
[illegible]	25	21	46	46	28	28	46.00	*	100	100	.97	.97	.00	.00			.00	.00	100.0	100.0
[illegible]	7	1					*	*			.00	.00	.00	.00			.00	.00	*	*
[illegible]	6	-3					*	*			.00	.00	.00	.00			.00	.00	*	*
[illegible]	115	62	145	62	104	80	64.73	-27.49	224	134	2.43	1.43	1.00	1.00	1229	-146	7.00	.00	16.5	28.0
[illegible]	95	-39	108	-66	104	-70	70.12	-17.31	154	-45	1.50	-.63	.00	.00	413	189	3.00	1.00	27.1	33.3
[illegible]	59	-13	52	-24	53	-24	60.22	-.40	38	-39	1.00	.00	.00	.00	336		2.00	.00	20.9	33.3
PVD	18	-1					*	*			.00	.00	1.00	.00			.00	.00	*	*
SCU	19	-2	2	2	1	1	66.66	*	3	3	.03	.03	.00	-1.00			.00	.00	*	*
RIC	15						*	*			.00	.00	.00	.00			.00	.00	*	*
SAT	2						*	*			.00	.00	.00	.00			.00	.00	*	*
SDF	12	-15	4	-72	2	-65	66.66	-19.70	6	-82	.07	-.93	.00	.00	89		1.00	.00	.0	.0
SFA	2						*	*			.00	.00	.00	.00			.00	.00	*	*
SJU	9	-6					*	*			.00	.00	.00	-1.00			.00	.00	*	*
SPO	31	31	137	-66	67	-38	51.11	-2.67	268	-105	2.93	-1.04	.00	.00	645	105	5.00	1.00	29.5	37.5
STL	56	56	50	-24	50	6	48.54	-37.50	103	17	1.00	.03	.00	.00	256	10	2.00	.00	26.6	33.3
[illegible]	78	78					*	*			.00	.00	1.00	1.00			.00	.00	*	*

Product Coordination

During the schedules planning process, the product coordination activity monitors the evolving schedule. During each iteration of the scheduling process, analyses are automatically produced which compare proposed levels of service to (Exhibit V) prior schedule periods, prior iterations of the proposed schedule, and prior and forecasted competitor service levels. When the schedule has been finalized, the reports are the vehicle for communicating knowledge of the product to advertising and sales.

MARKET PLANNING SYSTEMS

During the summer of 1974, the first phase of the marketing information system, a major data base management system, was implemented to support the planning function (Exhibit VI).

Components of the data base include:

- Eastern traffic and operational statistics from January, 1968, to the present. These statistics, in an audited form, are added to the data base on a monthly cycle. Onboard statistics (boardings, onboard passengers, seats, trips, etc.) are available daily throughout the month from the source file and are used for monitoring traffic flows.

- Competitor scheduled operations statistics from January, 1968, to the present. Schedule data are available from fifteen to twenty days prior to implementation and, at that time, are incorporated within the data base. Data reflecting irregular operations, such as a strike period, are subsequently deleted from the data base to minimize their

EXHIBIT V

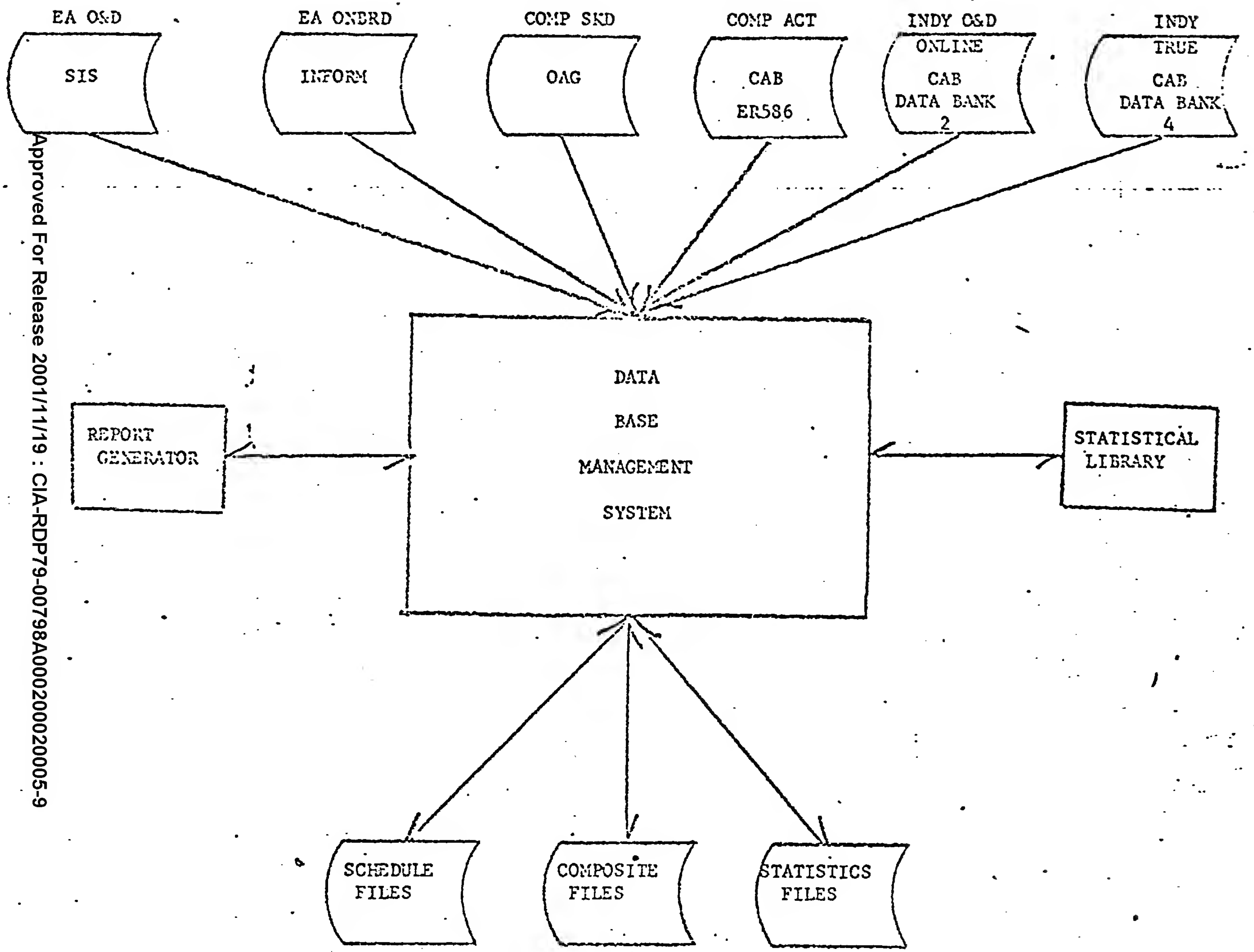
REGION OF ORIGIN ORIGIN CITY

SOUTH FLORIDA MIA

DES STA	EASTERN SCHEDULE SERVICE: 0-STOP TRIPS PLAN SKD	0-STOP TRIPS CHANGE PERCENT	1-STOP TRIPS PLAN SKD	1-STOP TRIPS CHANGE PERCENT	0-STOP SEATS PLAN SKD	0-STOP SEATS CHANGE PERCENT	IC AL CO	COMPETITORS ESTIMATED SCHEDULE: TRIPS 1-COMP 0-STOP	TRIPS TOTAL 0-STOP	TRIPS 1-COMP 1-STOP	TRIPS TOTAL 1-STOP	0-STOP SEATS FIRST COMP	SEATS TOTAL COMP	EASTERN: TRIP SHARE PLAN SKD	TRIP SHARE CHANGE ABS	SEAT SHARE PLAN SKD	SEAT SHARE CHANGE ABS	LOAD FACT 1975
AV[illegible]	.00	*	1.00	.00		*		.00	.00	.00	.00			*	*	*	*	
[illegible]	1.00	.00	1.00	*	89	.00	DL	1.00	1.00	.00	.00	90	90	50.0	.0	49.00	.00	75.[illegible]
[illegible]	1.00	.00	.00	*	134	.00	DL	.00	.00	1.00	1.00			100.0	.0	100.00	.00	44.[illegible]
[illegible]	2.00	100.00	1.00	25.00	210	100.00	DL	3.00	3.00	1.00	1.00	413	413	40.0	15.0	33.00	13.00	-6.[illegible]
[illegible]	.00	*	.00	-100.00		*	UA	1.00	1.00	1.00	1.00	98	98	.0	.0	.00	.00	
[illegible]	4.00	.00	.00	*	417	9.00	DL	.00	4.00	.00	.00		480	50.0	.0	46.00	2.00	76.[illegible]
[illegible]	3.00	-8.00	2.00	100.00	344	-33.00	DL	1.00	4.00	.00	1.00	106	574	42.0	-3.0	37.00	-10.00	65.[illegible]
JF[illegible]	4.00	-3.00	.00	*	722	-1.00	DL	3.00	9.06	.00	2.29	417	1675	30.0	-1.0	30.00	.00	71.[illegible]
LG[illegible]	5.24	-6.00	1.00	*	793	-6.00	DL	1.00	5.00	.00	1.00	106	846	53.0	-2.0	45.00	-2.00	74.[illegible]
[illegible]	4.00	33.00	.00	-100.00	449	33.00	DL	2.00	3.00	.00	.00	274	379	57.0	7.0	54.00	7.00	75.[illegible]
[illegible]	.00	*	1.00	.00		*	NA	.00	.00	1.00	1.00			*	*	*	*	
SY[illegible]	.00	*	1.00	.00		*	SO	.00	.00	.00	.00			*	*	*	*	
YU[illegible]	.00	*	1.00	.00		*	DL	.00	.00	.00	.00			*	*	*	*	
[illegible]	.00	*	.00	-100.00		*	DL	.24	1.84	.00	.00	117	252	.0	.0	.00	.00	
Y[illegible]	.13	*	.00	*	17	*	AC	1.00	1.00	.00	.00	135	135	11.0	11.0	11.00	11.00	
	24.97		10.00		3175			13.84	32.90	4.00	8.29	1760	4945					
TIX	.00	-100.00	.00	*		-100.00		.00	.00	.00	.00			*	*	*	*	2.[illegible]
	.00		.00					.00	.00	.00	.00							
ATL	6.00	.00	9.00	80.00	823	28.00	DL	6.00	10.50	2.55	2.55	1291	1848	37.0	.0	30.00	5.00	60.[illegible]
[illegible]	1.00	.00	.00	*	69	.00	SO	.00	.00	1.00	1.00			100.0	.0	100.00	.00	58.[illegible]
[illegible]	.00	*	1.00	.00		*	TW	.00	.00	.00	.00			*	*	*	*	
	7.00		10.00		912			6.00	10.00	3.55	3.55	1291	1848					
DF[illegible]	2.00	.00	.00	*	178	-3.00	DL	.00	4.00	.00	1.00		492	33.0	.0	26.00	-2.00	40.[illegible]
MS[illegible]	2.00	.00	.00	*	223	.00	NA	4.00	5.00	.00	.00	341	943	26.0	.0	19.00	.00	49.[illegible]
SAT	.00	*	1.00	.00		*	CO	.00	.00	2.00	2.00			*	*	*	*	
	4.00		1.00		401			4.00	9.00	2.00	3.00	841	1435					
CVG	.00	*	1.00	*		*	O	.00	.00	.00	.00			*	*	*	*	
I[illegible]	.00	*	1.00	-47.00		*	DL	.00	.00	2.00	2.00			*	*	*	*	
C[illegible]	.00	*	1.00	11.00		*	BN	.00	.00	.00	.00			*	*	*	*	
SDF	1.00	*	.00	-100.00	89	*	DL	.00	.00	1.00	1.00			100.0	*	100.00	*	
STL	2.00	3.00	1.00	.00	194	3.00	TW	1.00	1.00	2.00	2.00	123	123	66.0	1.0	61.00	1.00	45.00
	3.00		4.00		283			1.00	1.00	5.00	5.00	123	123					
	91.97		44.00		10967			46.84	116.26	26.55	47.26	7292	15517					

MARKETING INFORMATION SYSTEM

EXHIBIT VI



Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9

- Competitor actual operational statistics and onboard traffic for the period January, 1972-December, 1973. These data become available from twelve to twenty-three months after the fact, and are used to develop passenger departure time preferences.

- Competitor and industry traffic statistics from January, 1968 to December, 1974. These data are the true and online O&D traffic in each market. Market share for each competitor is available for online traffic. True origin and destination traffic is available for all domestic markets, and U.S. flag-carriers' traffic to international destinations. These data are added to the data base quarterly, but are six to nine months after the fact.

All data are maintained at the source level (flight or city pair). Additionally, a consolidated record containing frequently assessed statistics is maintained at the directional station pair level. Data base management software establishes linkages among related data at each level and across multiple levels.

A general purpose report generator provides the capability for generation:

- In variable formats
- With multiple data sets
- Up to five levels of summarization

Structured data files containing unique traffic statistics, including O&D passengers, onboard passengers, and enplanements are maintained as a subset to the data base. Interactive software components of the data base management system permits a library of statistical subroutines, including

- Regression analysis
- General purpose simulation system
- Time series analysis
- Trend analysis
- Variance analysis

These operate on this minidata base.

FORECASTING SYSTEMS

Two fundamental techniques are used to forecast traffic. One method uses:

- Stepwise regression with
- The Durbin Watson statistics to eliminate auto-correllated variables to develop the forecast algorithm

The other method uses:

- Time series analysis with
- Fourier analysis to overlay seasonality components

Economic indices used in the regression equation to forecast industry traffic include civilian employment, disposable personal income, government expenditure, corporate profits, and average air fare.

Forecasted industry data from each method are loaded into files, and an analysis of variance performed. Resolution of variances introduces the human factor into the forecast algorithm, which, functioning within an interactive environment (CRT interface with computer stored

files), can manually override the statistical projection with a judgmental forecast. Once the industry traffic has been forecast, regression analysis allocates the industry among all competitors by seat share, trip share, and market share. After the regression has been completed, a test for reasonableness is run to measure forecasted growth rates (Exhibit VII).

SIMULATION SYSTEMS

Although the simulation model was designed for use at the market and market group levels, it may also be used at the system level. Its primary function is to permit the analyst:

- To vary competitive assumptions and reforecast Eastern's position in the market

- To vary industry forecasts and recalculate Eastern's O&D traffic

- To vary both competitive assumptions and industry forecasts, and determine the impact on Eastern's O&D traffic

The analyst may elect to use either of two subsystems comprising the model. The first subsystem permits altering the independent variables in the algorithm by ± 1, forecast error value. The second subsystem requires that the analyst specify a numeric range for the variables. The model then uses a random number generator to vary the dependent variable over the specified range.

Another simulation tool is a general purpose simulator. The system was designed by IBM and is in general use by many airlines. This model is used primarily to compute a "risk factor" for each forecast. It permits the analyst to test assumptions relative to

MARKET RESEARCH

FS093C62

EXHIBIT VII

PAGE VII 6

VERSION NBR 01 DATE - 10/30/73

*** REASONABILITY ***

CITY PAIR REASONABLENESS TEST OF EASTERN SYSTEM FOR WINTER

RUN # 1 FORECAST 4 WINTER 1974

AIRLINE CODE	PC CHANGE FROM PREVIOUS YEAR — REVENUE PASSENGER MILES — ORG + DEST IND	COMPT	EAL	ON-CRD EAL	AVAILABLE SEAT MILES — IND	COMP	EAL	1974 EA SHARE PC — RPM	ASM	1973 EA SHARE PC — RPM	ASM	EAL PREMIUM/DEFICIT POINTS — 1973	1974	CHNG	LOAD FACTORS EAL CA-DFC — 1974	CHNG
AMERICAN	+21.0	+15.2	+4.0	+9.6	+3.5	+12.4	+3.1	+42.6	+40.4	+49.5	+42.4	+7.1	+2.2	-4.9	+61.2	+3.6
	RPMS = 1001.06 = 13.3 % OF SYSTEM RPMS							ASMS = 1605.23 = 11.4 % OF SYSTEM ASMS								
AIR CANADA	+19.9	-2.1	+32.3	+29.1	+3.3	-25.5	+40.1	+45.0	+59.4	+37.7	+53.6	-6.1	-14.3	-8.4	+52.5	-4.5
	RPMS = 54.67 = 0.7 % OF SYSTEM RPMS							ASMS = 104.47 = 0.7 % OF SYSTEM ASMS								
ALLEGHENY	-20.2	-52.2	-19.7	-17.1	-16.7	-13.6	-11.2	+179.4	+20.6	+164.8	+26.8	+134.6	+150.7	+12.7	+34.7	-2.5
	RPMS = 69.33 = 0.9 % OF SYSTEM RPMS							ASMS = 109.39 = 1.4 % OF SYSTEM ASMS								
AEROMEXICO	+5.6	-3.0	+12.3	+35.4	+24.0	+17.3	+31.7	+59.8	+49.8	+56.3	+46.9	+9.4	+10.0	+0.9	+65.4	+2.2
	RPMS = 159.91 = 2.1 % OF SYSTEM RPMS							ASMS = 344.43 = 2.4 % OF SYSTEM ASMS								
BRANIFF	+6.1	+25.0	-21.5	-20.9	+14.9	+22.6	-8.1	+37.4	+29.3	+51.4	+25.4	+26.6	+17.5	-9.0	+43.0	-6.5
	RPMS = 77.07 = 1.0 % OF SYSTEM RPMS							ASMS = 192.90 = 1.3 % OF SYSTEM ASMS								
DELTA	+5.5	+6.5	+4.6	+3.5	+11.4	+12.3	+7.2	+38.4	+36.7	+36.7	+37.8	+0.9	+1.7	+0.8	+51.0	-1.5
	RPMS = 2413.12 = 31.3 % OF SYSTEM RPMS							ASMS = 4729.54 = 32.6 % OF SYSTEM ASMS								
MOHAWK	+11.3	+6.9	+25.0	+5.7	+17.5	+6.0	+45.6	+30.3	+36.1	+32.3	+29.2	+3.2	+0.2	-3.0	+32.2	-11.8
	RPMS = 16.61 = 0.2 % OF SYSTEM RPMS							ASMS = 51.40 = 0.4 % OF SYSTEM ASMS								
NATIONAL	+2.8	+6.7	-1.3	-2.1	+10.4	+23.0	+11.8	+49.0	+38.5	+51.9	+47.6	+10.2	+10.5	+0.3	+48.1	-5.8
	RPMS = 1600.11 = 20.8 % OF SYSTEM RPMS							ASMS = 3326.65 = 23.0 % OF SYSTEM ASMS								
NORTHEAST	+2.2	+7.1	-0.3	-0.7	+29.2	+45.3	+10.2	+52.9	+38.7	+52.7	+43.4	+7.3	+12.2	+4.9	+51.3	-5.6
	RPMS = 1424.66 = 18.5 % OF SYSTEM RPMS							ASMS = 2776.35 = 19.2 % OF SYSTEM ASMS								
NORTHWEST	+5.3	-2.0	+26.8	+16.9	+10.2	+16.3	+14.7	+33.1	+27.3	+27.5	+27.6	-0.1	+5.4	+5.5	+52.2	+1.0
	RPMS = 563.62 = 7.3 % OF SYSTEM RPMS							ASMS = 1074.54 = 7.4 % OF SYSTEM ASMS								
PAN AM	+19.5	+15.9	-2.5	+15.3	+8.7	+12.1	+2.3	+36.5	+40.1	+44.8	+34.3	+2.5	-3.6	-6.1	+58.8	+5.3
	RPMS = 1024.23 = 13.3 % OF SYSTEM RPMS							ASMS = 1439.93 = 10.0 % OF SYSTEM ASMS								
PIEDMONT	+5.2	-32.3	+17.1	+11.4	+5.0	-16.0	+19.5	+85.5	+68.9	+76.8	+60.9	+15.9	+16.5	+0.6	+52.5	-3.8
	RPMS = 48.27 = 0.6 % OF SYSTEM RPMS							ASMS = 91.90 = 0.5 % OF SYSTEM ASMS								
SOUTHERN	+3.3	+18.1	-6.0	-5.0	+17.6	+33.6	+4.9	+55.4	+44.3	+43.8	+49.8	+11.1	+11.1	+0.0	+47.3	-4.9
	RPMS = 87.40 = 1.1 % OF SYSTEM RPMS							ASMS = 134.29 = 1.3 % OF SYSTEM ASMS								
TWA	+7.9	+11.0	+3.8	+3.9	-4.9	-16.7	+20.9	+39.4	+40.1	+41.0	+31.5	+9.5	-0.6	-10.1	+52.5	-7.8
	RPMS = 127.67 = 1.7 % OF SYSTEM RPMS							ASMS = 238.50 = 1.6 % OF SYSTEM ASMS								
UNITED	+3.8	+3.7	+8.8	+6.7	+10.4	+6.8	+15.6	+39.9	+40.6	+36.4	+32.7	-0.3	-1.1	-0.8	+53.6	-4.2
	RPMS = 1017.69 = 13.2 % OF SYSTEM RPMS							ASMS = 2011.78 = 13.9 % OF SYSTEM ASMS								

Eastern's various traffic components. For instance, if O&D traffic varied within a specific market group, the model will identify the impact of that variance on enplanements at all stations (Exhibit VIII).

SCHEDULE PLANNING

The schedule planning function has primary responsibility for the implementation of marketing strategies. Within the framework of the responsibility, schedule planning:

- Coordinates schedule-related decisions with station management to ensure that changes to service levels, equipment sizes, and departure times are integrated into the station planning process, and to provide input to the schedule building process relative to current competitive actions and local traffic conditions

- It also interacts with market planning during the schedule development process to provide product-related information for input to advertising and sales

The schedule planning process occurs in two phases: The first relates to the interface activities with market planning; the second phase with the iterative intuitive processes of schedule building.

Market Planning Interface

Marketing strategies reflect levels of service recommendations only, and are supported by market level passenger forecast data. Before they can be considered for implementation, they must be:

EXHIBIT VIII

MARKET RESEARCH VII
MARKET SCHEDULE PLANNING

AGUSTI-FRASER SAMPLE RISK ANALYSIS
FEBRUARY 1974 OPERATING PLAN STATION BOARDINGS

** PLAN MAJOR STATION,S BOARDINGS **
SYSTEM BOARDINGS - 1898

1973	OP PLAN 1974	MODIFIED 1974
302,204	320,628	286942
24,024	22,456	19867
12,684	12,404	10914
117,068	119,224	101476
55,216	58,184	54175
15,624	16,100	13688
20,804	2138,3	20370
12,936	13,748	12735
19,432	18,396	14430
7,028	7,448	6297
226,632	219,296	220646
18,732	18,508	14711
15,428	16,856	15738
340,508	357,112	310366

STA	1973	OP PLAN 1974	MODIFIED 1974
PBI	24,780	28,588	25769
PHL	53,760	56,840	51080
PIT	25,060	26,151	23538
RDU	27,720	31,388	28988
RIC	14,224	14,336	13271
SDF	14,196	14,336	12307
SJU	76,496	93,576	84465
SRQ	15,376	15,288	14332
STL	24,388	25,396	21879
STT	7,784	12,096	10447
STX	4,900	5,9368	5212
SYR	13,496	13,552	11822
TPA	55,160	53,116	48064
WAS	134,260	138,824	119228
YUL	13,644	14,420	11692

** MARKET GROUPS + MEAN FACTOR APPLIED TO EACH **

BAHAMAS	.85
BERMUDA	.87
CARRIBEAN	.84
COMMUTER	.81
EAST/WEST BUSINESS	.80
FLORIDA BUSINESS	.93
GULF	.78
JAMAICA	.85
MEXICO	.84
MIAMI/FT LAUDERDALE	.95
ORLANDO	.98
OTHER ATLANTA	.91
OTHER FLORIDA	.90
PIEDMONT	.93
SAN JUAN	.86
TAMPA	.87
VIRGIN ISLANDS	[illegible]

- Converted to flights with recommended
    - equipment
    - departure times
- Supported by flight level passenger forecasts including diversionary impact on other flights and markets
- Costed in terms of net new revenue and added expense

New flights which pass the test are integrated with existing flights and input to a Frequency Planning Model (FPM). Results of the FPM will determine both:

- The most profitable frequency of service in the lowest cost plane type which can fly each market
- Whether the capacity in the least cost plane type is sufficient to accommodate passenger demand

Several iterations of the model are usually required to optimize fleet mix and frequency of utilization by fleet type.

The second phase of the schedule planning process has been referred to as the "Computer-Assisted Iterative Intuitive Process." It is characterized by repetition of the human processes, as the schedule undergoes approximately thirty-six iterations during the final six weeks of development.

Output reports from each iteration are

- Schedule stability listing (Exhibit IX), which compares, at the flight level, the latest iteration with either
    - prior iteration or
    - prior schedule period
- And a one-stop flight listing (Exhibit X) which identifies through routings

EXHIBIT IX

TIME OF DAY/FLIGHT NUMBER/EQUIPMENT DETAIL REPORT

FAL75: DAY: 75/10/09
SUM75: CP-W DAY: 75/08/28
FAL74: DAY: 74/10/10

DATE: 05/05/75
PAGE 3

	FAL75							SUM75				FAL74			
MARKET	TIME	S F	NO	S F	EQ	S F	NIT	TIME	NO	EQ	NIT	TIME	NO	EQ	NIT
EWR-FLL								1130	745	83		1130	745	S3	
EWR-FLL	1230	X X	747		83										
EWR-FLL												1430	751	S3	
EWR-FLL	1830	X	759		S9			1830	759	S9					
EWR-FLL	2105		407		83	X		2105	407	83		2100	407	S9	
FLL-EWR								1030	742	S9		1030	742	83	
FLL-EWR	1330	X X	748		S9										
FLL-EWR								1530	752	83		1530	752	83	
FLL-EWR	1630	X X	756		S9										
FLL-EWR	2215		406		83			2215	406	83		2210	406	83	

Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9

 EXHIBIT X

MARKET RESEARCH
MARKET SCHEDULE PLANNING x
FALL 75 - VERS #8 - MON
***** DAB *****

MARKET	DIRCTN	ORG STA	TIME		TIME	DST STA	FLT NBR	STS	FREQ.	EFFEC DATE	DISC DATE	COMP ARVL
ATLDAB	ATLDAB	ATL	0034	JAX	0221	DAB	0438	105	MTWTFSS	750903	751201	0000
		ATL	1101		1207	DAB	0275	134	MTWTFSS	750903	751201	0000
		ATL	1355		1501	DAB	0585	134	MTWTFSS	750903	751201	0000
		ATL	1640		1745	DAB	0249	089	MTWTFSS	750903	751201	0000
		ATL	2010		2113	DAB	0131	105	MTWTFSS	750903	751201	0000
		ATL	2235	JAX	0024	DAB	0578	089	MTWTFSS	750903	751201	0000
	DABATL	DAB	0725		0837	ATL	0632	089	MTWTFSS	750903	751201	0000
		DAB	1008		1123	ATL	0284	134	MTWTFSS	750903	751031	0000
		DAB	1308		1420	ATL	0360	134	.....S.	750903	751201	0000
		DAB	1308		1420	ATL	0360	134	MTWTF.S	750903	751201	0000
		DAB	1600		1720	ATL	0244	134	MTWTFSS	750903	751201	0000
		DAB	2000		2114	ATL	0248	089	MTWTFSS	750903	751201	0000
		DAB	2200	JAX	2357	ATL	0466	105	MTWTFSS	750903	751201	0000
BALDAB	BALDAB	BAL	1741	ATL	2113	DAB	0131	105	MTWTFSS	750903	751201	0000
BNADAB	DABBNA	DAB	1008	ATL	1218	BNA	0284	134	MTWTFSS	750903	751031	0000
BUFDAB	DABBUF	DAB	0725	ATL	1114	BUF	0632	089	MTWTFSS	750903	751201	0000
CHIDAB	DABORD	DAB	1600	ATL	1916	ORD	0244	134	MTWTFSS	750903	751201	0000
		DAB	2000	ATL	2256	ORD	0248	089	MTWTFSS	750903	751201	0000
	ORDDAB	ORD	1302	ATL	1745	DAB	0249	089	MTWTFSS	750903	751201	0000
CLEDAB	CLEDAB	CLE	1553	CLT	1853	DAB	0719	089	MTWTFSS	750903	751201	0000
	DABCLE	DAB	1834	CLT	2120	CLE	0718	089	MTWTFSS	750903	751201	0000
CLTDAB	CLTDAB	CLT	1745		1853	DAB	0719	089	MTWTFSS	750903	751201	0000
	DABCLT	DAB	1308	ATL	1557	CLT	0360	134	.....S.	750903	751201	0000
		DAB	1308	ATL	1557	CLT	0360	134	MTWTF.S	750903	751201	0000
		DAB	1834		1942	CLT	0718	089	MTWTFSS	750903	751201	0000
DABJAX	DABJAX	DAB	0745		0815	JAX	0150	105	MTWTFSS	750903	751201	0000
		DAB	0836		0905	JAX	0678	089	MTWTFSS	750903	751201	0000
		DAB	1430		1459	JAX	0156	105	MTWTFSS	750903	751201	0000
		DAB	1505		1533	JAX	0666	089	MTWTFSS	750903	751201	0000
		DAB	2200		2229	JAX	0486	105	MTWTFSS	750903	751201	0000
	JAXDAB	JAX	0155		0221	DAB	0438	105	MTWTFSS	750903	751201	0000

MARKETING COMMENTS SHOULD BE DIRECTED TO ED FRASER (X6397,MIA)

The schedule planning objective during this iterative process is to maintain both

- preferred departure times for all flights in competitive markets
- through routings of potentially high load factor flights

SCHEDULE PLANNING SYSTEMS

Computer support for schedule planning is oriented to the basic activities of implementing and evaluating market strategies.

Frequency Planning Model

The evaluation function is supported by a Frequency Planning Model (FPM), designed to accomplish two main objectives:

- First, the allocation of flight services ("strings") to defined origin and destination (O&D) markets in an incremental manner, building an optimum schedule in such a way as to maximize the profitability of the airline
- Second, the evaluation of the effectiveness of a defined schedule by determining the net addition to profitability (incremental) of each flight "string" as well as determining the overall profitability of the entire schedule. Markets and segments with potential load factor problems are highlighted

The model considers round-trip flight "strings" such as NYC-BOS-CLT-MIA-CLT-BOS-NYC providing weighted desirability factors for non-stop, one-stop, multistop and connection services. Specific times of day are not considered although allowances are made for concurrent

operations on certain "strings" through a station for connection purposes. For each flight "string," the model utilizes the concept of a "basic aircraft." The "basic aircraft" is that aircraft which most easily pays for itself, or requires the lowest achieved load factor in order to return its cost of operation on that sequence.

In order to select optimum flight "strings" and the appropriate basic aircraft, inputs such as route license, market characteristics, aircraft operation characteristics, and individual station characteristics are considered by the frequency planning model. These inputs, along with miscellaneous others, are considered by the model in four phases: Input, Frequency, Capacity, and Output. The system re-evaluates each new flight pattern proposal in an iterative manner until, based on average yields and incurred operating expenses, return on investment is maximized.

A complete scheduling output is provided by flight segment, "strings," and total system. Of particular significance, either direct or calculated outputs include:

Frequency Planning Model Statistics

- Aircraft related statistics
  - Optimal fleet mix utilization
    - Fleet mix by ship type based on optimal schedule
    - Frequency of utilization by ship type
    - Alternative - Total fleet utilization (optimal schedule within "total" constraint)
  - Average stage length flown
    - Defined by ship type for each flight string
    - Summary by ship type--systemwide

- Available seat miles
  - Defined by ship type for each flight string
  - Summary by ship type--systemwide
- Revenue block-to-block (B/B) hours
  - Defined by ship type for each flight segment
  - Total B/B hours by ship type and segment
  - System total B/B hours by ship type
- Total departures by ship type, segment, and station

– Passenger related statistics

- Revenue passengers by flight segment
  - Revenue passengers by segment and ship type
  - Revenue passenger miles - flown (daily average, totals, and system summary)
- Revenue passenger movement
  - Enplanements by station and ship type
  - Deplanements by station and ship type with station activity summaries
  - Passenger originations-terminations-transfers by station
- Average fare structures
  - Average yield by flight segment
  - System summary--fare yields--all segments

- Passenger connections for complex terminals

- Market related statistics

  - Total O&D markets--all carriers

    - Total present passenger demand

    - Total projected passenger demand

    - Proportion of total airline market each O&D pair represents

  - Total O&D service--all carriers--each market

  - EAL share of O&D market--proportion of total market passengers

    - Projection of EAL market share

  - EAL share of O&D service--proportion of flights

  - EAL market share versus service share for each O&D city pair

  - O&D market demand load factor

  - EAL O&D achieved load factor

Time Sharing

The second function, implementing market strategies, is supported by time sharing programs on in-house IBM 360-65 computers. Time sharing programs operate within an interactive environment on a flight-level data base which contains:

- Historical eastern operational and traffic statistics
- Historical competiton schedule statistics
- Eastern's prior schedules
- Eastern's proposed schedules.

Output from the data base are "work sheet" reports which reflect the current version of the evolving schedule, and highlight variances to recommended changes relative to equipment size, departure time, and through service.

FLIGHT CREW ALLOCATION

The primary responsibility of the flight crew allocation function is to develop crew schedules which will support implementation of the general flight schedule while minimizing the cost of nonflying times. Involved in the crew scheduling process are the generation from the general flight schedule of round-trip flight routings called pairings; the integration of weekly and monthly exception flights (charters) into the paired schedule; and the combination of pairings into monthly allocations for bidding by crews.

For most United States airlines, there is an excess expenditure imposed upon the carrier for inefficient scheduling of its crews. In addition to paying for flying time, a carrier may be liable for additional pay guarantees for excessive on-duty time and excessive away-from-base time. The objective, then, of this function is to optimize the crew resource wherein excessive flight credit time will be minimum while complying with governmental and contract requirements.

FLIGHT CREW ALLOCATIONS SYSTEMS

Two computer systems support the crew allocation function:

- The interactive crew pairing system
- The trip pairing for airline crew scheduling (TPACS)

Crew Pairing

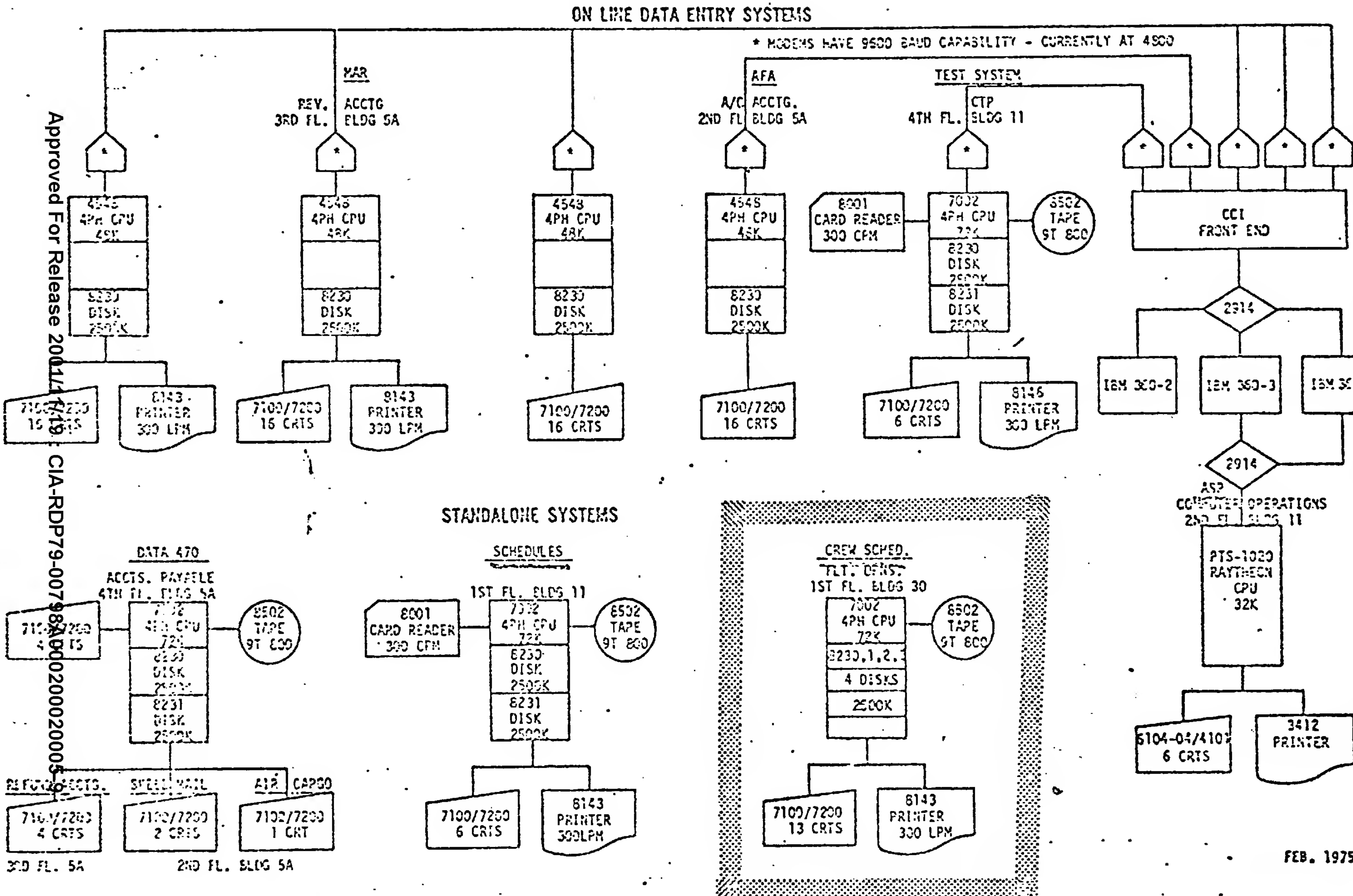
The crew pairing system (Exhibit XI) operates within a man-machine environment comprised to crew scheduling analysts and a small computer with four CRT input terminals. The four work stations operate independently; this permits simultaneous activity on four separate allocations, and provides an interactive capability allowing for the construction and modification of crew pairings, the determination of the legality and crew pay for any pairing, the evaluation of the impact of changes to the schedule, the display of credit and pay hour statistics by crew base, the production of hard copy listings of pairings, and the setup of special runs of TPACS which must be accomplished on the larger IMB-360 computers.

Functions supported by the system include:

- The general schedule load, which transfers a schedule of nonstop flight legs from tape to disk storage and establishes system time tables. This schedule, called a "master," is used by the bid schedule extraction for the creation of a bid schedule.
- The station table update, which allows the user to display interactively and update the tabled station parameters including Greenwich Mean Time adjustment factors and customs clearance codes.
- The bid schedule extraction, which creates a version of the schedule specifically intended for the pair construction activity by extracting a user-defined subset of a master schedule, and transfers it to the requester's desk pack, the transferred schedule being termed a "bid" schedule.

EXHIBIT XI

4 PHASE & RAYTHEON SYSTEMS



FEB. 1975

- The pair construction and update, providing the facility for interactive schedule display and update as well as pair construction and modification. Specific transactions generate displays of nonstop flight legs by flight number, or in the form of departures from a station. Updates to the schedule are accomplished by simply keying over the appropriate fields on the flight display. Schedule updates are automatically validated, and then directly introduced into the allocation. Pair construction requires the concurrent display of schedule information on one screen, and a pairing on another screen. Under transaction control, a segment is selected from the schedule display and subsequently inserted at any point in the pair display. Other transactions allow the deletion of segments from the pair and the insertion or deletion of duty breaks. Any change to the pair construction invokes an immediate recalculation and redisplay of pay and credit hours along with a complete legality check. The legality check returns diagnostic messages indicating the nature of any violations, and specifies the flight legs and/or duty periods at which they occurred. A legal pair is noted as such. Storing a legal pair to disk will update the base statistics which can be displayed at any time. A pair stored on disk in an incomplete or illegal status is called "flagged" and will be excluded from the base statistics.

- The pair reconciliation checks each pair of an existing allocation with a newly generated bid schedule, and performs any required updates to the pairings. The statistics of each

updated pair are removed from the base and flagged for later identification by the flagged pair display.

- The pair allocation load transfers an allocation from tape to the requester's disk storage, checking and updating each pair against the existing bid schedule. Pay computations and legality checks are performed and any illegal pairs are flagged. The base statistics are generated, using data from all the legal pairs.
- The pair allocation unload transfers an allocation from the requester's disk storage to tape for subsequent central site processing or to be used by the pair allocation load procedure.

Displays available to the analyst include such visual representations as

- Segments that are pairable for the allocation being constructed are identified by having two asterisks (**) in the first two positions of the record. Each segment displayed contains: flight number, type of equipment, departure time, arrival station, arrival time, effective and discontinue date, days operating, and pair number (if the segment is included in a pair). The segments are sequenced by departure time within effective date (Exhibit XII).
- A display by flight number that can be requested which will comprise all the flight legs for the requested flight number (Exhibit XIII).

EXHIBIT XII

STATION DEPARTURES DISPLAY

```
DEPARTURES FROM MIA
FLT #   EQ DEPT NXT ARIV FFF  DISC FREQNCY PAIR

** 0072 83 1210 ORD 1358 0501/0902 MTWTFSS  78
** 0364 S3 1220 ATL 1404 0501/0902 MTWTFSS  93
** 0898 B3 1225 RDU 1405 0501/0902 MTWTFSS  52
.. 0328 D9 1232 CLE 1500 0501/0902 MTWTFSS.
** 0008 B3 1236 EWR 1507 0501/0902 MTWTFSS  31
** 0953 B3 1245 SJU 1505 0501/0902 MTWTFSS  11
** 0036 B3 1250 PHL 1514 0501/0902 MTWTFSS  40
** 0042 B3 1252 BOS 1540 0501/0630 MTWTFSS  93
** 0182 S3 1255 BDL 1534 0501/0902 MTWTFSS  61
.. 0190 S9 1255 DCA 1510 0501/0902 MTWTFSS
** 0018 B3 1300 JFK 1536 0501/0630 MTWTFSS  70
** 0283 B3 1316 FLL 1335 0501/0902 MTWTFSS   8
.. 0851 L' 1320 NAS 1410 0501/0902 MTWTFSS
.. 0539 S9 1323 FPO 1355 0501/0902 MTWTFSS
.. 0690 S9 1334 STL 1506 0501/0902 MTWTFSS
** 0993 B3 1335 STT 1554 0501/0902 MTWTFSS  36
.. 0524 S9 1355 MSY 1446 0501/0902 MTWTFSS
** 0020 B3 1400 LGA 1642 0501/0902 MTWTFSS  78
                  CONTINUED

   DA,MIA                         APR25 16:19:15
```

EXHIBIT XIII

FLIGHT DISPLAY

```
FLIGHT # 0993
AL EQ    DEPT   N    ARIV   N  EFF DISC   FREQ   PAR

** B3 ORD 0900, MIA 1250, 0501/0902 MTWTFSS 78
** B3 MIA 1335, STT 1554, 0501/0902 MTWTFSS 36
** B3 STT 1625, SJU 1650, 0501/0902 MTWTFSS 36

  F0993                        APR25 16:21:18
```

- A pair construction display that includes (for each leg in the sequence) flight number, station codes, local departure and arrival times, flying times, and layover time. A duty break is indicated by the presence of block-time total, scheduled on-duty time and duty tour accumulation. Sequence totals are shown last, and include: away time, away credit, block time, and tour credit (Exhibit XIV).
- A display of base statistics which will present crew block and pay hours by equipment type for each domicile. The base statistics are always updated to reflect the current status of the allocation (Exhibit XV). And, finally,
- A display that shows which stations have operations that are not yet included in any pairings (Exhibit XVI).

TPACS

TPACS is used by the crew pairing function to optimize the crew scheduling resource. Input to the model is a paired flight schedule generated within the interactive crew pairing system. Output from the system is an optimized paired schedule, wherein excess flight credit has been minimized yet governmental and contractual constraints have been satisfied.

Processing Description

One phase of the crew allocation may be formulated as a programming problem. Consider the matrix [a(i,j)] where the rows represent flight segments and the columns indicate possible legal pairings:

EXHIBIT XIV

PAIR DISPLAY

```
      EFF: 0501      PAIR 036
  10% CREDIT                              BLOCK L/O   DTY  TUR
 B3 0993 MIA STT 1335 1554 219                   31
 B3 0993 STT SJU 1625 1650  25                  145
 B3 0975 SJU BGI 1835 1956 121                   30
 B3 0975 BGI POS 2026 2110  44   449  1055     850  503
 B3 0774 POS FDF 0805 0848  43                   32
 B3 0774 FDF ANU 0920 0956  36                   27
 B3 0774 ANU STX 1023 1100  37                   28
 B3 0774 STX SJU 1128 1200  32                  140
 B3 0778 SJU SDQ 1340 1424  44                   50
 B3 0737 SDQ SJU 1514 1600  46   358  1445C    925  523
 B3 0760 SJU SDQ 0645 0727  42                   48
 B3 0763 SDQ SJU 0815 0900  45                  130
 B3 0998 SJU STT 1030 1051  21                   39
 B3 0998 STT STX 1130 1149  19                   41
 B3 0998 STX MIA 1230 1459 229   436               929  527
AWAY TM:  5039 CR:  1550        1323               *1553

   P A I R   L E G A L

                                   APR25 16:24:31
```

BASE STATISTICS DISPLAY

** BASE STATISTICS **

BASE	HOURS	B3	LE	D9	L1	SH
ATL	BLOCK	282				
	PAY	327				
BOS	BLOCK	67				
	PAY	74				
DCA	BLOCK	68				
	PAY	78				
JFK	BLOCK	257				
	PAY	280				
MIA	BLOCK	283				
	PAY	328				
ORD	BLOCK	55				
	PAY	57				
SJU	BLOCK					
	PAY					
*SYS	BLOCK	1013				
	PAY	1144				
	CREDIT	131				
	%	13				

B APR25 15:34:37

$a(i,j) = 1$ - when flight i is covered by pairing j

$a(i,j) = \emptyset$ - when flight is not covered by pairing j

$c(j) =$ - cost of pairing j

$x(j) = 1$ - when pairing is a part of the solution

$x(j) = \emptyset$ - if pairing not selected in the solution

The SUM $[a(i,j)] = 1$ (summed over j) expresses the fact that each flight must be covered only once. The problem then becomes one of minimizing SUM $[x(j)c(j)]$. A potential difficulty with this approach relates to the concept of generating all possible pairings. Fortunately, this difficulty is easily avoided within the current environment.

Rather than addressing the total schedule, TPACS assumes the role of a scheduling analyst who usually works with thirty-to-forty flight legs at any one time. The problem, then, is reduced to optimizing a "set" of the schedule. To further aid the model in this activity, "sets" are selected from an existing solution (manual) with high cost pairing considered first. For each "set" or subproblem, the model will run to completion or until a time limit is exceeded. Legality is maintained throughout the process. A geographical test is also applied to minimize reprocessing of subproblems containing geographically isolated pairings.

SCHEDULE DEVELOPMENT

The flight schedule is a key determinant of the airlines' potential for profit and its exposure to costs. Small changes in the schedule can have a dramatic impact on the number of flight crews needed; the peak station facilitates and manpower requirements; on on-time performance, hence the ability to maximize connecting revenue

passengers; and the overall competitive posture in the market.

Interrelationships among these factors are subtle and complex, particularly when considered in relation to Eastern's dense route network. Schedule development, then, is the focal point of the resource allocation process. To be responsive to the diverse demands placed upon a schedule, it is imperative that updates to a schedule under development and reports reflecting the updated schedule occur within a real time environment. We believe that we at Eastern have achieved that prerequisite.

At the beginning of the schedule development cycle, available resources are "known" and include:

- Aircraft availability by fleet type
- Crew capability (total hours by crew base)
- Station resources, including
    - Available gates
    - Manpower
    - Overnight accommodations
    - Ground equipment

Also available are:

- Operational guidelines such as
    - Ground time requirements for through, turnaround, and connecting flights
    - Maintenance schedules
    - Overnight requirements
- Aircraft utilization by fleet type

The objective of the scheduling process is to create the "best of all possible worlds" wherein the utilization of available aircraft is maximized through optimization of flight crew capability, station manpower resources, and ground facilities. Optimization occurs during a period of approximately six weeks. During this period, daily iterations of the scheduling process occur: between iterations, changes to the schedule include recommended departure times and aircraft rotations.

Changes are reported to cognizant functional areas via a flight listing and a computer-readable file. Recommended changes from each functional area are then fed back to the scheduling group and input into the next iteration.

SCHEDULE DEVELOPMENT SYSTEMS

Schedule development is provided three different types of computer support:

- Conventional batch processing, which produces hard copy output for all participants in the schedule development process
- Time sharing systems on in-house IBM 360-65 computers, with the capability to perform analytical studies during the development process
- And an online schedule update system using minicomputers with CRT input devices and a high speed printer to support limited output requirements. This system is the key to the iterative processes which are prerequisite to optimization. The primary output from the system is a flight schedule stored on magnetic tape which feeds both the batch scheduling system and the crew pairing system.

Online Schedule Update System

The online schedule update system is designed to accommodate two phases of schedule development:

- Future schedules, which are event-oriented and consider annualized departure times in relation to a station, with limited effort devoted to ordering flights by rotation sequences during this period
- Current schedules which take the event-oriented schedule and add a flight orientation

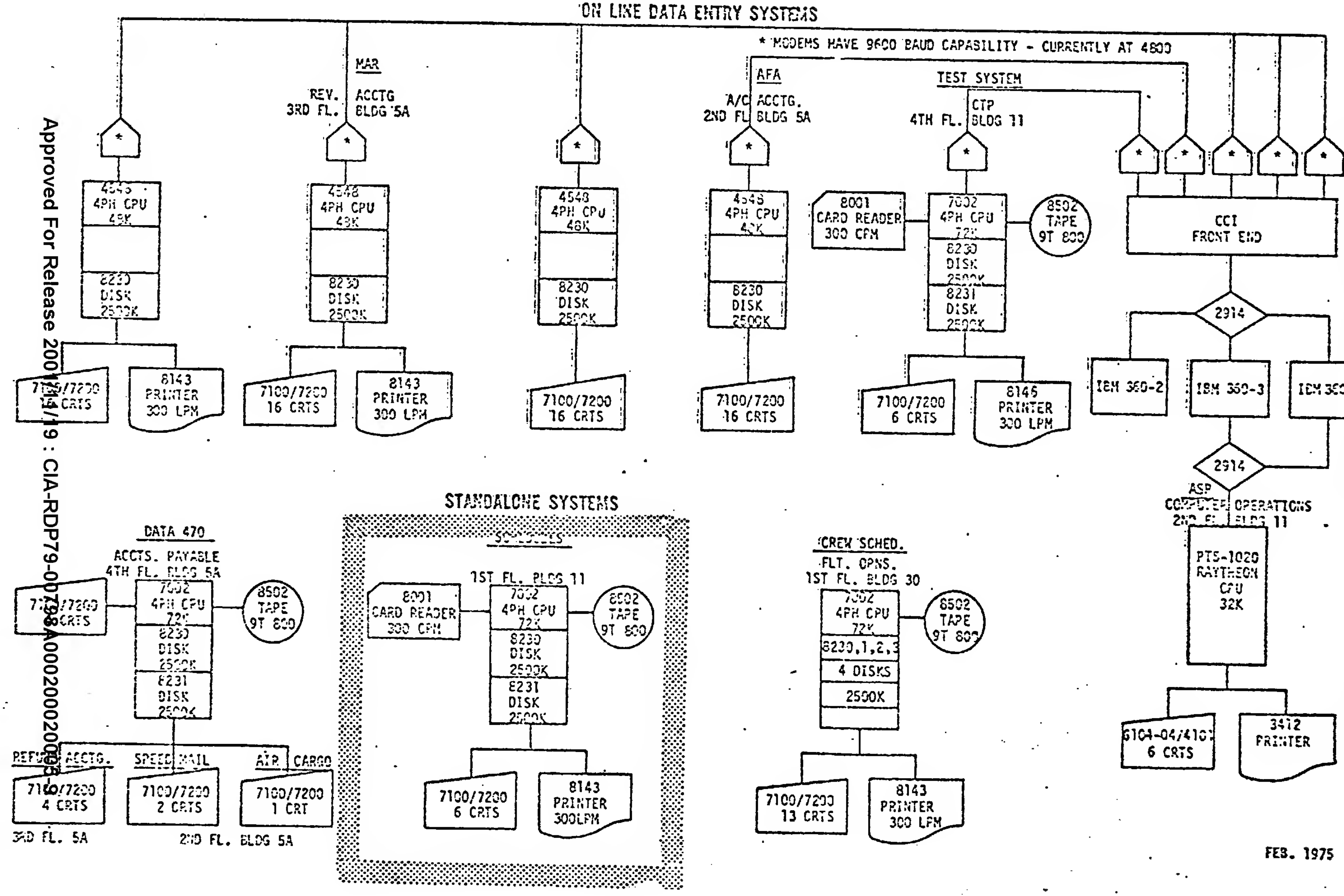
Since both the future and current schedules functions are working concurrently on different schedule periods, system design considered an online CPU-sharing capability. To achieve the design objective, a work station configuration of three CRTs and a keyboard is provided for each functional system. A printer is shared between the two work stations. Of the three CRT's, one is active, while the other two are used for display of reference information. Each system has its own displays, formats, and transaction codes, peculiar to its processing requirement. Full screen cursor control is provided, and data elements to be altered are simply overwritten. Use of function keys simplifies keyboard operation and triggers routine functions such as screen printing.

Computer and Peripheral Hardware

The hardware supporting this application (Exhibit XVII) is the System IV/70 Model B of Four-Phase Systems, Inc. It is a 72K byte mini-CPU with 2.0 usec cycle time. The configuration includes the CPU, an 800 bip tape drive, a 400 lpm printer, and two work areas,

EXHIBIT XVII

4 PHASE & RAYTHEON SYSTEMS



FEB. 1975

each having a 2.5 million byte removable disk with 90 msec average access.

For this application, the Four-Phase computer operates entirely alone. Communication with central site IBM 360-65 application is via magnetic tape. The software consists of a control subsystem and two application subsystems. The control subsystem contains all common subroutines required for:

- Data conversion and manipulation
- Disk and tape I/O processing
- The application of subsystem control: specifically, the assignment of time segments

Each application subsystem is allocated to a separate segment of memory and resides on a separate disk pack with the data base it will process. It includes all transaction processing peculiar to its own activities. As a transaction is entered at one of the keyboards, the resident control program enters the control module for the requested transaction and passes over the control. The control module then accesses all other modules needed to complete the transaction and, where necessary, structuring overlay sequences.

Each system provides a set of transactions to display and update the schedule data base and provide schedule summary statistics. Customized keyboard/cursor control and function keys provide simplified keyboard operation and initiate screen printing.

Display Capability

In a flight-oriented system, displays most frequently used include

- Flight summary
- Flight itinerary (Exhibit XVIII)

  Flight summary display of Flight 180 which has two itineraries, both effective January 31, 1973, and discontinued on April 28, 1973; Flight 180 terminates in Washington except on Saturday when it continues to Philadelphia

  Flight itinerary display of itinerary A from display above; display of turn from/to flight numbers, classes of service, and meal codes has been inhibited

  The same itinerary is displayed, with only classes of service and meal codes inhibited. (Note that Flight 180 turns from Flight 183 in Tampa and turns to Flight 577 in Washington)

  The same itinerary, with all data elements displayed, including classes of service and meal codes

and - Schedule Statistics Summary (Exhibit XIX).

  Display of schedule statistics summary for February, 1973

  Selected statistics are available seat miles (ASMs) and aircraft block hours

  In each category, the first figure is the computed total, while the second figure is the deviation from a planned level input separately at load time for each category

  Display of schedule statistics for March

FLIGHT SUMMARY AND ITINERARY DISPLAYS

```
FLT 0180                                   VERSION 1601
 A 730131 730428 MTWTF.S B3 TPA DCA
                                              XVIII
 B 730131 730428 .....S. B3 TPA DCA PHL
```

```
0180.A EFF 730131  DIS 730428 EQ B3 FQ MTWTF.S

 FROM ....   .   :TPA  1235  ......  ...  .  ..
 .  ....  .  :1428  DCA  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...      TO:....    .
             A
```

```
0180.A EFF 730131  DIS 730428 EQ B3 FQ MTWTF.S

 FROM 0183   .   :TPA  1235  ......  ...  .  ..
 .  ....  .  :1428  DCA  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...      TO:0577    .
```

```
0180.A EFF 730131  DIS 730428 EQ B3 FQ MTWTF.S

 FROM 0183   .   :TPA  1235  FYL...  LL.  .  ..
 .  ....  .  :1428  DCA  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...  ....  ......  ...  .  ..
 .  ....  .  :....  ...      TO:0577    .
 .  ....  .  :....  ...
```


EXHIBIT XIX

SCHEDULE STATISTICS SUMMARY

730201 730228 FEBRUARY 1973 XIX VER 1601

TOTALS

	ASMS (000)		BLOCK HOURS	
L1	369,945	46,648-	3,963	31
D8	25,543	64,707-	427	1,159-
AC	0	0	0	0
S8	203,034	58,061	2,513	702
AK	0	0	22	22
BN	2,380	0	56	1-
B3	568,191	13,392	14,893	410
AB	0	0	24	24
S3	383,653	108,360-	7,764	2,075-
AJ	0	0	35	35
L3	139,998	139,998	2,644	2,644
QC	0	0	2,024	5
QP	224,021	2,773	5,839	21
AH	0	0	47	47
D9	64,334	1,968	2,876	86
AD	0	3	36	36
S9	502,018	21,177	16,666	620
AG	0	0	176	176
SH	61,363	0	2,663	17
C9	0	0	0	0

CTD

730301 730331 MARCH 1973 VER 1601

AVERAGES

	ASMS (000)		A/C MLS	
L1	13,821	2,201-	61,155	1,925-
D8	912	2,315-	6,553	16,652-
S8	7,201	2,157	36,854	10,844
BN	85	0	664	0
B3	20,379	520	207,949	5,310
S3	13,698	3,778-	103,769	28,629-
L3	4,993	4,993	37,822	37,822
QC	0	0	33,787	61
QP	7,917	456	80,791	4,757
R7	47,672	2,201	461,782	10,322
D9	2,296	69	34,786	1,647
S9	17,704	292	201,151	3,322
SH	2,184	0	20,414	0
C9	0	0	0	0
R9	22,184	361	256,382	4,378
LT	379	152-	4,357	1,743-
LS	37	0	450	0
CE	[illegible]	[illegible]	[illegible]	[illegible]
RE	416	152-	4,787	1,743-
RV	91,087	52	327,523	14,216

In this display, aircraft miles were selected in place of block hours

In addition, the display of daily average numbers was requested in place of totals for the period

In an event (station) oriented system, the primary display is an event summary by station (Exhibit XX).

- Display of departures from Atlanta. In this display, the selection criteria was for stretch DC9 departures after 1800 hours.

Audit Protection

A key feature of the system is immediate audit/validation of all updates. The audit provides "turn" validation, verifies flying time against a master table, and checks for duplicate operations of the same flight number on the same day. Station, equipment, meal, and class of service codes are verified against internal tables. Error conditions generate immediate diagnostics which must be corrected before the flight can be stored in the schedule data base.

Transactions

Clusters of transactions perform the fundamental functions of retrieving flight data from the disk and formatting displays, supporting and auditing update activity, and altering the disk resident data base.

EXHIBIT XX

STATION DEPARTURES DISPLAY

```
STATION DEPARTURES    ATL    VERSION: 1-0143-0000
     O DEPT FLT  NXT FREQNCY EFF    DISC N TURN EQ A
XX
 :  .. 1815 0326 BUF MTWTFSS 0131/0428 . 0000 S9 .
 :  .. 1815 0786 MSP MTWTFSS 0131/0428 . 0000 S9 .
 :  .. 1815 0786 MKE MTWTFSS 0131/0428 . 0000 S9 .
 :  .. 1820 01[illegible] LGA MTWTFSS 0131/0428 . 0000 S9 .
 :  .. 1822 0616 RIC ......S. 0131/0428 . 0000 S9 .
 :  .. 1822 0616 RIC MTWTF.S 0131/0428 . 0000 S9 .
 :  .. 1825 0546 PVD MTWTFSS 0131/0428 . 0000 S9 .
 :  .. 1830 0142 SYR MTWTFSS 0131/0428 . 0000 S9 .
 :  .. 1830 0244 ORD .....SS 0141/0428 . 0000 S9 .
 :  .. 1835 0125 EWR MTWTFSS 0131/0428 . 0000 S9 .
 :  .. 1835 0136 DCA .....S. 0131/0428 . 0000 S9 .
 :  .. 1835 0324 CLT MTWTFSS 0131/0428 . 0000 S9 .
 :  .. 1835 0565 GSP .....S. 0131/0428 . 0000 S9 .
 :  .. 1835 0566 GSP MTWTF.S 0131/0428 . 0000 S9 .
 :  .. 1900 0137 IND MTWTFSS 0131/0428 . 0000 S9 .
 :  .. 1950 0131 MOB MTWTFSS 0131/0428 . 0000 S9 .
 :  .. 1955 0345 CSG MTWTFSS 0131/0428 . 0000 S9 .
 :  .. 1955 0567 MIA MTWTFSS 0131/0428 . 0000 S9 .
                   CONTINUED
```

A single operation might require the execution of an entire series of transactions. For example, to update a flight record on the disk, the analyst must request a display of the flight and one or more audits of update attempts. Once the flight has been validated, he has the option of replacing the flight on disk or erasing the update and restoring the flight to its original form. Additional transaction sets provide the ability to update various internal reference tables (e.g., master flying time, city pair mileage, etc.) as well as generating displays of combinations of equipment as summary information.

Data Base Load/Unload

Each system has the ability to produce, as well as receive, the entire schedule data base and related files in the form of an IBM 360-65 compatible tape, each again in its own format. A card deck specifying system parameters is read concurrently with the load tape. During the load process, the flight or event records are subjected to a partial audit, and any invalid information is flagged and indicated on a printout.

Following a period of schedule development, the updated schedule file, together with the reference files specified by a control card deck; are written to tape. This tape serves two purposes—as input to the IBM 360-65 batch processes and as a backup checkpoint.

CONCLUSION

This paper has had two objectives; namely,

- To provide an overview of the major functional areas within the resource allocation process
- To review solution systems implemented to facilitate resource optimization

The techniques and methodologies incorporated into our systems represent a significant breakthrough in relationship to the predominantly manual environment of the past, but these are not the ultimate answers.

As with other airlines, we continue to experience difficulty in matching capacity to market demand. This problem will persist until such time as we can read the minds of our competitors and the traveling public.

A Software System for Urban Transportation Planning

Robert B. Dial *

The Last 20 Years

Sometime in the 1950s a certain Hudson Valley transportation planner, a direct descendant and namesake of Rip Van Winkle, dozes off over his drawing board to sleep the traditional twenty years. While the latter-day Van Winkle dreams his unimodal dreams, undisturbed by social and environmental nightmares, unaware of energy crises, his more lively colleagues slave away.

Upon awakening in the 1970s, he sleepily looks out his office window and immediately notices the apparent ineptitude of his old colleagues. It is obvious that in the planning and improvements he slept through, Americans invested trillions of dollars in automobiles, roads, parking facilities, traffic signals, policemen, traffic courts, hospitals, insurance companies, tire factories, oil industries, drive-ins, and billboards--all in deference to the private automobile/highway system. Yet, in spite of this enormous capital expenditure, traffic problems were no closer to being solved, and the cities smelled awful and looked worse than they did twenty years ago.

Coming to their own defense, his fellow planners argue that they should not be faulted for the current state of affairs. They were misguided in their ignorance of the issues. No one urged them to consider costs and benefits except those supporting the popular demand for more cars and more roads. They lacked both the technical and fiscal wherewithal to plan, much less build, anything but the automobile-dominated existence we now suffer at enormous national expense.

* Director of Planning Methodology and Technical Support Division, Urban Mass Transportation Administration

It was not until the 1960s that federal legislation admitted that urban man could not move by car alone and looked to transit for help. Like an aged football player abruptly recalled from retirement to substitute for the limping superstar, public transportation was dusted off, given an aspirin, and sent into the dying seconds of the game. Renamed "Mass Transportation" (perhaps to connote the movement of slugs rather than people), it was ordered to reduce congestion by "them" so that "our" automobiles might go faster. With less than one percent of the capital budget spent on the automobile/highway system, it was asked to solve the problems the automobile could not as well as those the automobile had caused. And, to make matters worse, it was given no federal subsidy for operation.

After Van Winkle gets the sleep out of his eyes, he dutifully checks the records for the twenty years of his nap. Something of a student of human nature and of history, and now with the freshest mind in the business, he is neither surprised nor alarmed by the misconceptions or the misdirections. He regrets the waste of time, money, and talent, but he understands it in terms of our political economy and is not disposed to judge it harshly.

Van Winkle is most disappointed to see that our planning procedures differ so little from those he was using when he dozed off. He saw that the technical expertise needed to solve problems (problems unknown in the fifties) had increased by an order of magnitude. Immediately, it was clear to him that he would need new methods to deal with ideas that he had never heard of: priority lanes, road pricing, dial-a-ride, PRT, environmental impact statements, energy conservation, and quality of life. The problems are new, and the ground rules for

their solutions have changed, but Rip notes that our present day tool kits hold the same tools that had rusted to pieces in his battered old box as he slept away twenty long, long years.

Since Van Winkle refuses to give up despite the staggering problems of urban transportation, the least we can do is help him replace his tools, find new tools for new work. And as we awaken with him into a new era of transportation planning, his clear view of the stunning differences between the fifties and the seventies can help us decide what kinds of tools are needed.

Lessons Learned

Certainly there are four lessons that Van Winkle's experience teaches us, and for us to learn them with him is a condition of future success in transportation planning--an absolute necessity if our urban transportation systems are to be saved from inexorable decay.

The first lesson is that the transportation problem can be solved only at the local level. It is apparent that the problem was made worse by a federal tilt toward highways during the last twenty years, and federal policy that earmarks dollars for specific modes, regardless of local needs and desires, aggravates rather than ameliorates the situation. Any effective solution will likely require a better use of the automobile coupled with vastly improved public transportation.

The second lesson is that we must better exploit the transportation resources that we have and not automatically assume, in response to a problem, that what we need is more. Our superb highway system is fifty percent underused about ninety percent of the time. Too often, roads are conceived of as providing for the movement of cars and trucks, not of people and goods, while, in fact, at certain times it is

advantageous to ban cars and trucks from some segments of the road system. Public transportation vehicles, pedestrians, and cyclists should receive much higher priority in the planner's mind and on the city streets.

The third lesson is that urban transportation planning, implementation, and operation must be coordinated without an aritificial administrative and jurisdictional partitioning of functions and responsibilities. Planners must guide builders. Operators must trust planners. Planners must be informed by builders and operators. In the past, these people scarcely knew one another. Today they must work together.

The fourth lesson is that the planner must consider a much larger set of options and issues. He must look for more and better transportation alternatives, and the evaluations of these alternatives will, in large measure, be based on nontransportation issues. Not only is today's problem more acute, but the constraints on feasible solutions are tighter. More technical expertise is required.

Today, the planner must plan a system, not merely design appendages to growing freeways. He now must justify his recommendations with lengthy alternative analyses, examining vastly different and sometimes radical proposals. He must describe and defend the numerous potential impacts of a proposed plan to vociferous politicians and a suspicious public, whose questions are selfish, diverse, and microscopic. A decision to build will never again be based on a simplistic travel time measure. Many other criteria, often conflicting, must be addressed.

We must now let our brave old friend, our newly awakened planner, go back to his work. But for one last time we should acknowledge our debt to the Van Winklean perspective as we begin to describe UMTA's

current concerns and goals. That historical perspective has been very important to us, as UMTA research and development have begun to reflect the recent federal awareness of the very different planning problems of a new era with new complexities.

Needed: Improved Planning Tools

Planning techniques now in common use are slow and costly. Slow because they use a hunt-and-peck system to find a good plan. Costly because of long turnaround times and high data costs. Their most serious weakness is their inability to evaluate multimodal planning alternatives accurately and responsively. At best, they plan effectively for one mode, the private automobile.

Local planners are keenly aware of these shortcomings. They must respond quickly to local policy questions. Despite their inadequate resources, they must go ahead and plan with what they have. Piecemeal efforts of local planning agencies to improve tools often cost more than their marginal success is worth. The federal government's research and development of improved planning techniques will be especially valuable and welcome at the local level.

UMTA's Responses

For as many years as large computers have been available, state and local agencies have used them to plan. UMTA R&D best helps local planners by packaging for their use the best research and development products in computer software. In this way UMTA can require local planning to improve and provide the technical and fiscal support for that improvement.

Accordingly, in 1972, the UMTA Office of Research and Development began an R&D program to:

- Research and develop improved planning techniques
- Implement these techniques in generalized computer software
- Pilot test software in urban areas to ensure its appropriateness and demonstrate its utility
- Distribute the software to local planners
- Provide technical backup by training users and responding to queries from the field

The result of this program will be the Urban Transportation Planning System (UTPS). UTPS is a package of computer programs for site-specific planning of multimodel transportation systems. The package is evolutionary, being constantly enlarged and updated. Its ultimate goal is a streamlined, easy-to-use set of modular tools applicable to several planning activities.

UTPS PLANNING PHILOSOPHY

Two considerations affect the design of UTPS. First, variations in local budgets, data, and issues bring about many different planning situations, and no one model fits them all. Second, to be easy to use and yet adequately sophisticated, its technical complexity must in large measure be invisible to the user--like that of a telephone.

To accommodate the variety of planning situations, UTPS distinguishes three overlapping, sequential, and iterative planning activities: long-range planning, short-range planning, and system surveillance.

Long-Range Planning

There are two types of long-range planning. One searches for a strategy, and the other articulates in some detail a design within a selected strategy. We call these Sketch Planning and Tactical Planning.

Sketch Planning is the preliminary screening of possible multimodal configurations or concepts under varying assumptions regarding alternative futures. (It has both manual and computerized versions.) Using highly aggregated measures, it compares a large number of proposed policies in analytical detail just sufficient to support strategic decisions. Especially needed in long-range regional planning (ten-twenty years), sketch planning, at minimum data costs, yields preliminary estimates of multimodal network's capital and operating costs, patronage, wide corridor traffic flows (by mode), service levels, and land development implications. It also estimates such factors as energy consumption and air pollution. It compares all these data with those available about other networks and provides the information needed for broad policy decisions.

The demands on such a strategic model for long-range planning are challenging. First, it must be very easy to use and quick to evaluate an alternative strategy. Far-future options are limitless. Scores of them must be considered, and thus each must be done quickly. Second, the model must be designed to be able to simulate the performance of modes which are as yet unspecified. Third, it must deal explicitly with uncertainty. Two of the most vexing uncertainties are those associated with the socioeconomic and land developments, and those associated with the costs and performance of new transportation technologies.

Sketch planning input is characterized by a small (less than 800 nodes) but rich abstraction of an (abstract) multimodal network in which supply-demand equilibria are explicitly considered. Trip generation, distribution, modal split, and assignment--traditionally four different technical steps--are handled in a single step. Outputs are immediately relevant to the issues. Sketch planning is an aggregate, multivariate system evaluator and comparer. It evaluates a single system alternative at less than ten percent of the cost of existing long-range planning techniques.

The planner remains in the sketch planning mode until he completes his comparisons of possibilities or finds a strategic plan worthy of consideration at the tactical level.

Tactical Planning treats the kind of detail appropriate to midrange (five to ten years) planning and identifies the best configuration within a given strategic concept uncovered in the sketch planning phase. The input and analytical techniques are close to those of today's state-of-the-art regional and corridor planning studies. Inputs include the location of principal highway facilities and delineated transit routes. These feed a network model that addresses any automobile/transit vehicle interaction. Disaggregate demand forecasting techniques are applicable.

The cost of examining an alternative in midrange planning is ten to twenty times its cost in sketch planning; although "default" models, which assume away certain data requirements, might be run for an inexpensive first look. As contrasted with sketch planning, tactical planning can provide disaggregated cost and benefit measures related more accurately to the citizens affected.

At this level of analysis, the outputs are estimates of transit fleet size and operating requirements for specific service areas, refined cost and patronage forecasts, and level-of-service measures for specific geographical areas and, where necessary, a program for staged implementation. Household displacements, noise, localized pollution, and aesthetic factors can also be evaluated. Apparently promising plans can be analyzed in yet further detail, and problems uncovered at this stage may suggest a return to sketch planning to accommodate new restraints.

Short-Range Planning

The development of good short-range planning tools brings the greatest return for the model development dollar. This is especially true because the strong tradition of pure highway planning, pre-occupied with long-range, capital-intensive programs, is little help in the evaluation of immediate action programs. Short-range planning brings the tools and analytical techniques badly needed to optimize the use of a city's existing transportation resources. The development of these tools has high priority at UMTA.

As in long-range planning, there are two distinct types of short-range planning activities. One is a quick evaluation of broad, area-wide, transportation strategies, and the other is the preparation of a detailed delineation of an "optimal" system design reflecting a given strategy. In the former, the difference from long-range strategic planning is that the short-range case requires more accurate cost/benefit estimates. Fortunately, greatly improved accuracy is obtainable. By comparison, the feasible transportation options here are very limited and the costs and capabilities of individual system

components are accurately known. And, in the short-term, human behavior and demand for transportation are less difficult to forecast. Accordingly, a much more precise evaluation is possible. Some examples of the kinds of policies a short-range strategic model can address are:

- Area-wide dial-a-ride service
- Widespread designation of automobile fee zones
- Road user tax or increased gas tax
- Order of magnitude increase in transit fleet size or exclusive guideway (lanes)
- Broad changes in parking policy

Detailed delineation of the plan and the system's expected costs and benefits is required prior to a final decision to implement. The outputs of long-range tactical planning models and the short-range strategic models are usually too abstract for engineering design purposes, but, as the time to implement projects draws near (five days to five years), detailed simulations can be made to refine design parameters. Some examples of activities at this stage are:

- Detailed evaluation of the extension, rescheduling, or repricing of existing bus service
- Simulation of bus priority lanes or signal systems
- Analysis of passenger and vehicle flows through a transportation terminal or activity center
- Comparison of possible routing and shuttling strategies for a demand-activated system

Analysis at this detailed level can be prohibitively expensive except for subsystems whose implementation is very likely, and in which

such design refinements bring substantial increases in service or significant reductions of cost or uncertainty. It is effective in near-term planning only when the large number of exogenous variables can be accurately observed or estimated.

Surveillance

Knowledge of the current state of affairs is a prerequisite to any planning. It is essential that existing highway and transit systems and their users be monitored to ascertain the service provided, to whom, and at what cost. Such data are needed for supply and demand model verification and calibration as well as system evaluation. In addition to the traditional traffic counting, user-oriented surveys of such things as convenience and travel time must also be maintained. Information on the citizens' travel patterns and socioeconomic attributes are also needed.

Besides permitting a continual scrutiny of transportation services, costs, and usage, the data from surveillance supports near-term planning to eliminate problems, such as overloaded links, inadequate transportation opportunity, and better resource utilization.

UTPS FUNCTIONAL CHARACTERISTICS

To support the planner in the four stages identified above, UTPS acts as a highly interactive system, using time-shared computers with on-line graphics terminals--vastly different from the present slow-motion, error-prone batch operation. Interactive browsing through network and land-use data, both digital and graphic, speed up the planner's evaluations. Maps, charts, and graphs replace the millions of numbers that now overwhelm him. Graphic input via an electronic tablet speeds his data entry and run setup. An interactive network

design model allows him to specify or to modify his plan virtually instantaneously. Many analytical processes are run while the planner waits at his CRT, giving him "instant turnaround." To guarantee successful execution, longer analyses which require batch processing are "dry run" interactively before submission. Later, the planner interactively browses through the outputs of the batch process.

The UTPS program library includes data management routines, graphics routines, statistical and mathematical programming packages, and specific planning models and algorithms--the software needed to examine transportation supply and demand at each of the three planning levels described above. UTPS modules meet uniform software design standards, and adherence to those standards allows UTPS to add new software and provide improved analytical techniques as they become available.

Among the most important modules are those for system evaluation, demand estimation, network aggregation, data acquisition, and data management.

System Evaluation

The system evaluation tool is an open-ended set of reports and graphics, selected for the use of local planners, who may include their own reports on local issues. UMTA adds new reports as national issues arise. Local planners can compare significantly different network conceptions and make detailed analyses of the minor peturbations of a given network. They can evaluate present and proposed systems according to current and future demands. The other modules described below also directly support system evaluation.

Demand Estimation

Planners making demand estimates may choose from three kinds of models: off-the-shelf default models for local use without site-specific parameter estimates, default models with locally calibrated parameters, and user-made models that can be integrated with an existing module with little programming effort.

Algorithms for establishing supply-demand equilibria provide the capacity to determine route and mode selection equilibrium, origin-destination demand equilibrium, and land development-transportation equilibrium. The software supports the development, calibration, and application of both aggregate and disaggregate models.

Network Aggregation Models

Among the improved tools under research are the network aggregation models, useful at all levels of planning. The automatic reduction in size of the coded network description speeds up the computing process by providing the data base most efficient for an analysis. There are three network aggregation techniques: subarea windowing, region-wide abstraction, and subarea focusing.

Subarea windowing is the most straightforward technique. There is software that physically extracts a subarea of the network and collapses external demand to within the subarea's periphery. It can be used for detailed analysis and short-range planning when external demands are assumed to be fixed.

Region-wide abstraction is technically more difficult. The computer reduces detailed networks to a specified level of abstraction by aggregating links, nodes, and zonal data, yielding a network amenable to sketch planning. This permits movement from the tactical

or short-range stage back to the sketch planning stage, thus allowing rapid macroscopic evaluations of detailed networks.

Subarea focusing is the most difficult technique because it combines windowing and abstraction. A subarea of interest is windowed, but the links outside the window are not deleted but abstracted, so any modification of the subarea's internal network can have the appropriate effect on external demand. This is accomplished by increasing network abstraction as distance from the window increases. Subarea focusing greatly improves the effectiveness of traditional long-range (tactical) planning--reducing its cost and increasing its accuracy.

Data Acquisition

While data collection is essential to planning in general and system surveillance in particular, the notoriously large sums of money spent for data acquisition should be channeled into more productive analyses. Planners need more efficient data gathering techniques. UTPS must couple modern sampling techniques with the power of an on-line, time-shared computer and modern data entry hardware to speed the collection, editing, **correcting,** and to reduce the cost of survey data. Also, a disaggregate travel demand data base is available to researchers to eliminate the need for more data in certain cases. Detailed network coding manuals show the planner the quickest way to input his transportation system's characteristics.

Data Management

The data management system is used to specify network and land-use configurations, edit data, evaluate systems, etc. A good data management system must allow the planner to execute programs and

interact with the data base without detailed knowledge of the data base's design. It should also be possible to provide a common source of data for all UTPS modules, allow efficient data base modifications, avoid a proliferation of data files, and furnish a repository for output from computational modules.

Besides the many computational similarities (e.g., matrix manipulation), there are also many common data requirements among the three levels of planning analysis, such as network descriptions, land-use data, and graphic data. Therefore, data preparation time and user training time are reduced, and the software is fully exploited. At any time the user may modify the basic network or land-use data by using the interactive network design program. The modifications can be additions, deletions, or the updating of any or all elements: but the basic integrity ofthe original design and its predecessors is preserved in a tree-like file structure. At any time, the planner may analyze any version of the network. In UTPS a single data base might contain scores of networks, all quickly available for analysis.

The planner can design his network while describing it to the computer in a natural, graphic fashion. He sits at a cathode ray tube and, using a stylus or light pen, draws the network. He draws it either by explicitly entering nodes, links, transit lines, etc., or alternatively by circumscribing geographical areas of homogeneous service which is described parametrically (e.g., street spacing, number of bus stops, etc.).

The UTPS package can generate maps, charts, or graphs. When the software processes a request for graphics, it saves the results in the graphics file of the data base. The file contains the points, lines, and annotations that constitute the graphic in a standard format.

The planner may browse through the available graphics at any time, recalling, combining, modifying, or displaying those needed, without the expense of regeneration. Attribute or land use data can be overlaid on network plots and the graphic directed to a display tube or hardcopy plotter.

CONCLUSION

All components and capabilities described above are among the current future objective of the UTPS development effort. All are presently in a research or development stage. A few initial products have already been released to the planning community. Most are scheduled for future delivery.

In its present, skeletal state, UTPS consists of thirteen software modules and attendant documentation forming a fairly powerful suite of programs which run in the batch mode on the IBM 360/370 series of computers. Basically comprising a traditional transportation model, it best supports long-range tactical planning, but can be forced into limited service for the strategic or short-range planner. It includes highway and transit network analysis models, demand forecasting models, matrix manipulation, and limited graphics capabilities. It installs easily at the user's computer facility and is being continually improved with new developments.

It is hoped that within three years, UTPS will evolve to include all the capabilities discussed above. It will be in a form which allows it to be fairly readily installed on non-IBM computers and will exploit minicomputer and nationwide computer network technologies. The result will be a ubiquitously available software system, which should be an aid to federal, state, and local planners who search for effective solutions to today's complex and moving transportation problems.

ACKNOWLEDGEMENTS

The design and development of UTPS is the product of individuals too numerous to list. Among the institutions they represent are: Barton-Aschman Associates; Cambridge Systematics; Comnet Corporation; Consad Research Corporation; Creighton-Hamburg, Incorporated; Deleuw, Cather and Company; DTM Incorporated; First Data Corporation; Massachusetts Institute of Technology; Metropolitan Washington Council of Governments; National Bureau of Standards; the MITRE Corporation; Peat, Marwick, Mitchell and Company; Planning Research Corporation; R. H. Pratt Associates; Wilbur Smith and Associates; TRW Systems Group; Alan M. Voorhees and Associates; and the Urban Mass Transportation Administration, which provided financial support for the operation of the above institutions as well as the preparation of this paper and disclaims any Federal policy which might be read into it.

AN EVALUATION OF THE AIR QUALITY IMPACTS OF TRANSPORTATION CONTROL POLICIES IN U.S. URBAN AREAS

Gregory K. Ingram *

1. INTRODUCTION

Over the past two decades in the United States concern about the quality of the environment increased dramatically. Concurrently, the air quality in many urban areas deteriorated, and scientists demonstrated direct linkages between pollutants such as photochemical oxidants and the exhaust emissions of motor vehicles. These scientific findings and the worsening urban air quality led to the 1970 amendments to the Clean Air Act which set ambient air quality standards and empowered the U.S. federal government to require new motor vehicles to reduce their emissions of certain pollutants by approximately ninety percent. Although the phased reduction of motor vehicle emission rates was the major policy proposed to meet the ambient air quality standards, the 1970 legislation also allowed urban areas to limit the extent of pollutant generating activity if the reduction in emission rates proved insufficient. In many urban areas of the U.S., it appears that transport activity will have to be curtailed somewhat if the ambient standards are to be achieved on schedule. A variety of transportation control policies including improvements in public transit, better traffic controls, and restrictions on automobile use have been proposed to curtail transport activity in U.S. urban areas.

This paper analyzes a number of these transportation control policies in two U. S. cities using a computer-based simulation model. The objective of the analysis is to clarify the potential role of transportation control policies in improving air quality and to rank

*Harvard University & National Bureau of Economic Research

alternative policies in terms of their cost and effectiveness. The computer model that is used also could be applied to other urban transportation policy problems, and the structure of the model is briefly summarized before the transportation control policies are presented.

2. THE TASSIM MODEL

The analysis of transportation control policies is based on the Transportation and Air Shed Simulation Model (TASSIM), developed with funds from the U.S. National Science Foundation and the U.S. Department of Transportation.[1] This model forecasts levels of travel activity, pollutant emissions, and pollutant concentrations for distinct zones within an urbanized region. The overall model, as illustrated in figure 1, consists of three major components: a transportation submodel, an emission submodel, and an air diffusion submodel. Each of these submodels incorporates existing modeling techniques. The transportation submodel is based upon the urban transportation planning model used extensively in transportation studies. The emissions submodel employs emission factors published by the U.S. Environmental Protection Agency to calculate mobile emissions. The diffusion submodel was developed by meteorologists.

The transportation component of the TASSIM model is an adaptation of the standard urban transportation planning (UTP) approach.[2] The transportation submodel is composed of the usual steps of trip generation, trip distribution, modal split, and network assignment. In the trip generation step, the total number of daily person trips which originate in, or are destined to, each zone is forecast as a function of the household, land use, and employment characteristics of the zone.

FIGURE 1

MAJOR COMPONENTS OF

THE TRANSPORT AND AIR SHED SIMULATION MODEL

TRANSPORTATION SUBMODEL

Forecast distributions of transportation activity (trips and speeds) in each zone by mode and purpose of trip.

EMISSION SUBMODEL

Derive area source emissions from automobile travel using speed/emission factors. Input stationary area source emissions by zone and large point source emissions by location.

DIFFUSION SUBMODEL

Generate pollution concentrations in each zone based on emissions and meteorological factors.

These trip forecasts can be envisioned as the row and column totals from a matrix that displays trips by origin-destination pair. In the trip distribution step, a matrix of trips by origin and distribution is created that is consistent with the trip generation totals. The trip matrix is produced by employing a distribution rule such as a gravity model or an intervening opportunities model. In the mode split step the trip matrix is subdivided into two or more trip matrices indexed by mode (automobile, transit, etc.). Mode split predictions are typically based on the socioeconomic characteristics of the population and the relative cost and performance characteristics of the modes. Finally, in the network assignment step, the trips from each origin zone to each destination zone are loaded on a spatially detailed representation of the transportation system. The network assignment typically assigns trips to the shortest path in the network.

Three major changes distinguish TASSIM's transportation submodel from the standard UTP approach. First, the transport submodel has been calibrated at a level of detail which is more aggregate than is usual for such models. The Boston version of TASSIM, for example, considers 122 zones and 582 interzonal transport links for both auto and transit. The TASSIM links comprise a spider network whose links represent either all of the streets and highways or all of the transit routes that connect zone nodes. Thus, the model produces link loadings on aggregates of highways and arterials rather than flows on individual transportation facilities. The more aggregate nature of the TASSIM model produces enormous computational savings. Moreover, the spatial detail retained is sufficient to represent the distribution of transport activity among zones and the changes in that distribution caused by transportation control strategies.

Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9

The second change in the transportation submodel involves the prediction of transit ridership which is shared by the modal split and network assignment steps of the model. In the standard UTP model, auto trips and transit trips are assigned separately to their respective networks. TASSIM employs a more flexible representation of network assignment and mode choice. The two categories of trips represented are transit trips that are made completely on the transit mode, and other trips that begin as auto trips. These latter trips can either switch to the transit mode at some point or else continue to to their destination on the auto mode. The use of a composite network (containing highway links, transit links, and intermodal links) for trip assignment enables the TASSIM model to simulate the effect of parking charges and other vehicle restraint schemes on auto trips in a more realistic manner than would be possible with the standard UTP model. The auto networks used in standard models are often so large (containing 10,000 to 30,000 links) that it would be prohibitively expensive to expand them to composite networks, whereas the composite network for the TASSIM model is still of a reasonable size.

Finally, the third change incorporated in the transport submodel is its use of a diversional routing or multipath assignment procedure in the network assignment step instead of the all-or-nothing or capacity-constrained assignment procedures. The probabilistic multipath assignment procedure used in the network assignment step selects a set of feasible paths from each origin node to each destination node and distributes the trips to these paths by means of a weighting scheme.

The spatial traffic patterns forecast by the transport submodel permit the calculation of auto vehicle miles travelled per zone as well as auto vehicle cold starts per zone.[3] Average vehicle emission rates for a given fleet are multiplied by the estimates of vehicle miles travelled and cold starts to yield emissions by zone from mobile sources. Emissions from small stationary sources, such as home heating units, are added to mobile emissions to give average area-wide emissions for each zone. An area source diffusion model then predicts the zonal concentrations of each pollutant which result from the emission patterns and the meterological characteristics of the urban area.[4] Emissions from large stationary sources, such as electric generating plants, are treated as individual point sources, and a second diffusion model is used to forecast the pollutant concentrations in each zone that stem from these large individual emitters.[5]

These diffusion models both assume that the relation between emissions and concentrations is linear and additive, and that the pollutants represented are nonreactive. That is, the models assume that emissions can be related to concentrations by an equation of the form:

$$C(I) = A(I,J) * E(J),$$

where $C(I)$ is the n-dimensional row vector of the concentrations in zone $I (1 \leq I \leq n)$, $E(J)$ is the n-dimensional column vector of emissions from zone $J (1 \leq J \leq n)$ and $A(I,J)$ is the n x n transformation matrix.

Two automotive pollutants, HC (hydrocarbons) and NOX (oxides of nitrogen) combine in the presence of sunlight to form photochemical oxidants, which are important secondary pollutants in many metropolitan areas. The reaction processes that produce photochemical oxidants are not well known, and at this time there are no simple diffusion models

that can predict spatial concentrations of oxidants with an acceptable degree of accuracy. Therefore, photo-chemical oxidant concentrations are not forecast in this analysis, and results are obtained only for the oxidant precursors, HC and NOX. The diffusion models used in this analysis are elementary, even primative by some standards: but, unlike proportional rollback models, they represent the effects on air quality of redistributing emissions over space. If transport controls are applied to selected portions of an urban area such as the central business district, existing patterns of transportation activity and mobile emissions will be altered. Thus the evaluation of transportation controls requires a diffusion model.

A major reason for using a model to analyze air quality is the consistent framework it provides for systematically evaluating policies that reduce or redistribute emissions in urban areas. When the TASSIM model is calibrated to a particular city, a wide range of transport control strategies can be considered within the same land use and transportation environment. The choice of the study site is of crucial importance: analyzing an atypical city may result in unwarranted conclusions about policy impacts.

Rather than searching for the "average" urban area, this study has selected two cities, Boston and Los Angeles, which represent quite different types of urban development in the United States. Table 1 presents several socioeconomic and transport statistics that summarize the characteristics of these two cities. Figures 2 and 3 display the analysis zones used in each city. Los Angeles has three times as many people, but four times as much area and four times as many automobiles as Boston. Moreover, in a typical weekday, residents

Table 1

SUMMARY STATISTICS FOR BOSTON AND LOS ANGELES FOR 1970

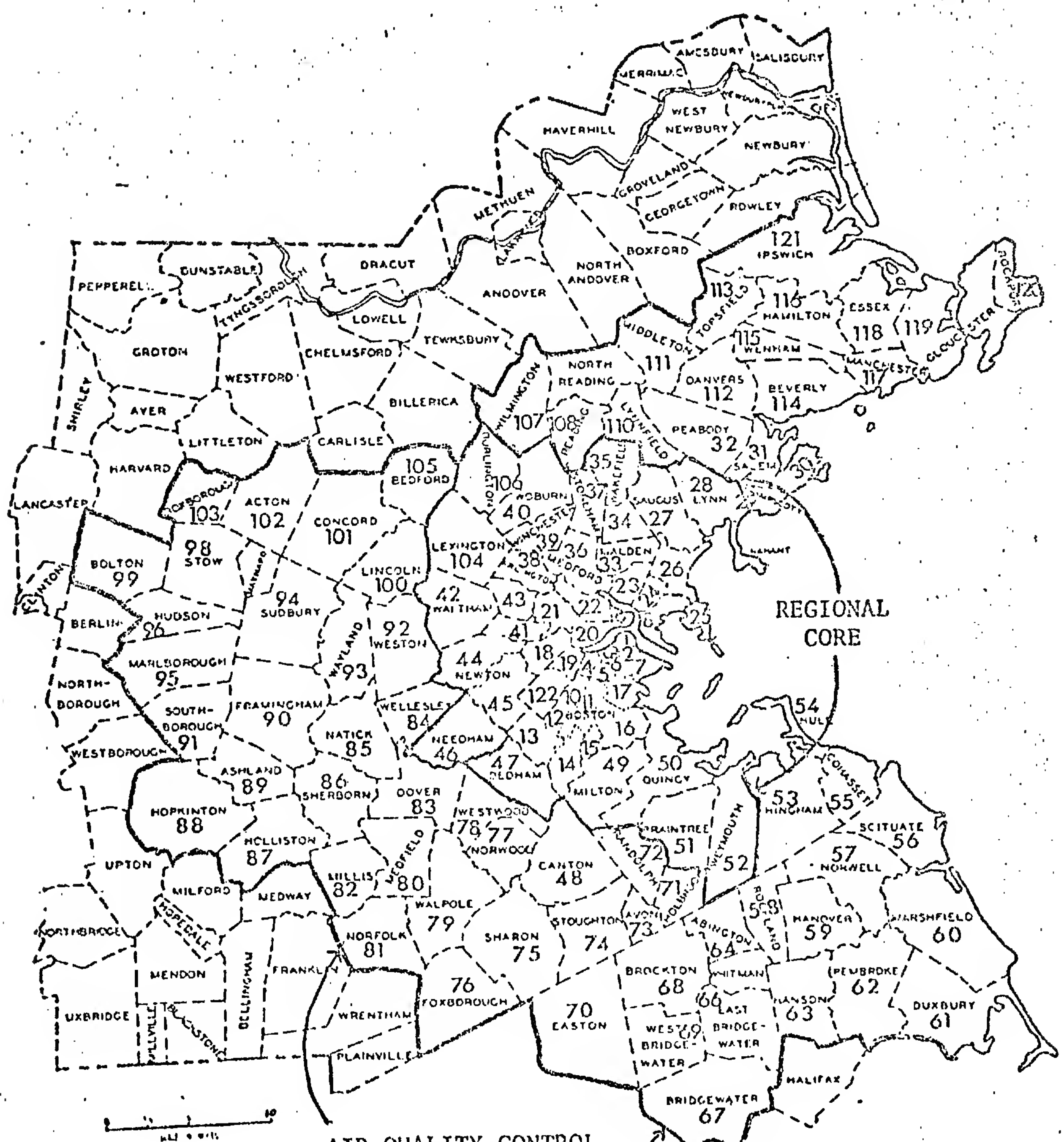
Item	Boston	Los Angeles
Population	3,023,000	9,008,400
Automobiles Owned	1,096,000	3,930,200
Area (sq. mile)	1,400	5,285
Daily Person Round Trips	3,125,038	12,243,975
Trips to Central Business District	384,774	416,281
Transit Originating Round Trips	348,252	190,052
Transit Originating Share of Trips	11.1%	1.6%
Daily Vehicle Miles Travelled	21,960,336	165,936,096
Average Auto Round Trip Length (Miles)	9.91	19.13
Average Transit Round Trip Length (Miles)	8.54	14.17
Average Auto Speed (mph)	19.13	36.46
Hours of Travel	2,040,852	6,450,413
Average Daily Hours of Travel per Person	0.68	0.72
Emissions by Pollutant in Grams/Second		
CO	19,090	64,215
HC	2,839	17,601
NOX	1,057	9,541

Source of transportation and emission statistics: TASSIM simulations and LARTS Base Year Report.

Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9

Figure 2

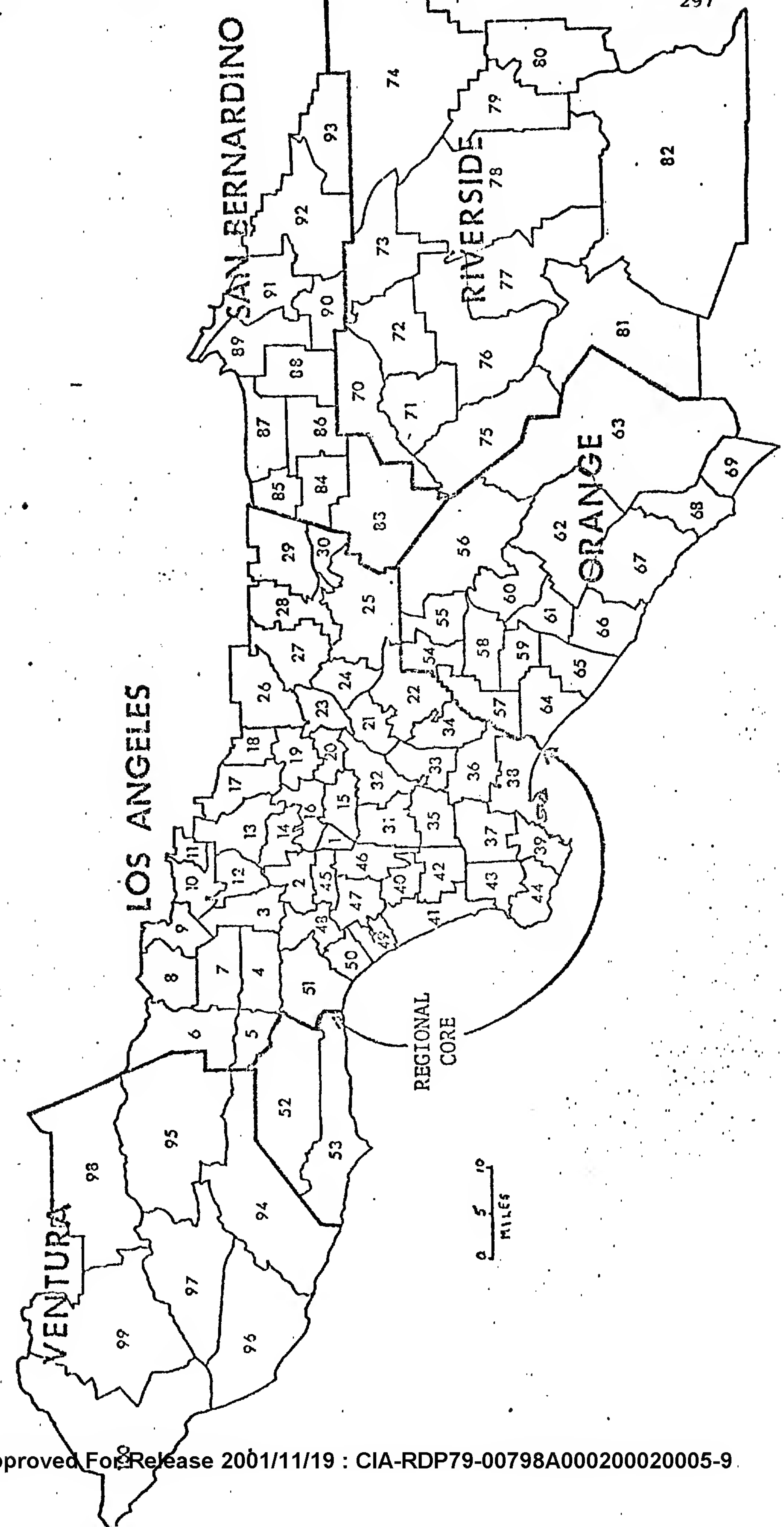
BOSTON ZONAL SYSTEM



AIR QUALITY CONTROL
REGION BOUNDARY

Figure 3

LOS ANGELES ZONAL SYSTEM



of Los Angeles make four times as many trips as the residents of Boston. The number of trips destined for the CBD area, by contrast, is nearly identical for the two cities. A much higher share of Boston's trips are made by public transit, and this fact, coupled with the differences in trip level and trip lengths, produces a total figure for daily auto vehicle miles travelled in Los Angeles that is eight times as large as the figure in Boston.

These statistics reflect fundamental differences in urban structure between Boston and Los Angeles. The former is typical of older eastern U.S. cities that developed with a strong orientation toward a central core and that possess well-developed public transport systems. The latter is typical of the newer auto-oriented cities of the West and South that have more diffuse land use patterns and that have transport systems more reliant on the automobile. Studying these two cities using the TASSIM model should indicate how differences in transportation and land use patterns alter the efficacy of transport controls.

In many respects the model versions of Boston and Los Angeles that have been developed can be envisaged as small-scale models of the two metropolitan areas. Various policy "experiments" are carried out with the two urban areas for a particular base year--in this case 1970--for both cities. This base year calibration provides a point of departure or benchmark for the analysis of transportation control or emission reducing policies. To execute policy experiments certain model parameters are changed to represent a particular policy- and the model is operated. The outputs from the model with the policy in place are then compared with the outputs from the benchmark case

so that differences in pollutant levels and transportation activity can be identified.

Some of the model outputs can be used to estimate the area-wide costs of a policy. For example, if a particular policy increases (or reduces) total vehicle miles travelled or the total number of hours that people spend travelling each day, these incremental resource costs (or savings) need to be added to the administrative costs of implementing the policy.

Several effectiveness measures are also produced by the model. One is a prediction of the number of annual region-wide person-exposures to pollutant levels that exceed the official ambient air quality standards. For each pollutant and each zone, the model first calculates the number of times per year that the ambient standards are exceeded.[6] The annual number of violations in the zone is then multiplied by the population of the zone to produce the number of person-times of exposure for the zone. Finally, this number is summed over all zones and pollutants to produce a regional exposure index. To summarize this procedure:

$$REX = \sum_{K} \sum_{I} SV(K,I) \cdot POP(I)$$

where REX = regional exposure index;

K = pollutant types:

SV = standard violations per year for particular pollutant

POP = zone population:

I = individual analysis zones.

In addition to region-wide summary statistics such as vehicle miles travelled, the number of daily trips made by transit, and the regional exposure index, both the Boston and Los Angeles version of the model produce transportation, emissions, and air quality data by zone that are only imperfectly reflected in the region-wide statistics. It should be emphasized, moreover, that the spatial disaggregation of the TASSIM model is its most important feature, since many policies being recommended for improving air quality are localized ones. For example, suggestions to increase parking charges, restrict vehicle access, bypass through traffic, and augment traffic flow capacity usually are envisaged for particular areas such as central business districts or other portions of central cities. Since many of these policies may be expensive to implement, it is important to be able to predict the spatial distribution of the improvements in air quality that they will sustain and the disruptions they will generate. The level of spatial disaggregation in the TASSIM model permits the analysis of air quality improvement policies at a scale that is consistent with the areas that are being considered for the application of transportation control strategies.

3. IMPACTS OF POLICIES ANALYZED IN BOSTON AND LOS ANGELES

Transportation control strategies can be separated into two main categories: those policies that can be applied relatively easily with fairly well-known impacts and costs, and those policies that are more difficult to implement and whose impacts and costs are less certain. Of the policies discussed in this section, the locally applied, or target area schemes, tend to fit in the first category, whereas the region-wide plans typically belong to the second category.

The emission reducing and transportation control policies listed in table 2 have been analyzed with the Boston and Los Angeles versions of the TASSIM model. Although most of the policies were applied to both urban areas, a few policies simulated for Boston could not be replicated for Los Angeles. It was not possible to consider the land-use policies of centralization and decentralization with the Los Angeles model because of its method of predicting interzonal flows of trips. In addition, testing a rapid rail extension is an option viable only in Boston, which has an existing system. Both the aggregate and spatially detailed information from each policy run were studied to analyze each policy. The brief discussions which follow summarize the conclusions of these efforts.

a. Reducing Mobile Source Emissions

Reducing new car emission rates by ninety percent over a five-year period is a principal method for improving air quality in urban areas set forth in the 1970 amendments to the Federal Clean Air Act.[7] The lower emission rates of new autos will reduce average automotive emission rates as new, low-emission vehicles gradually replace the existing fleet of high-emission vehicles in a metropolitan area.

The TASSIM model uses average emission rates which are calculated outside the model, so simulating the effects of reduced vehicular emissions over time simply requires substituting altered emission rates. Trip-making patterns are held constant. When a future year is simulated, the 1976 and 1980 emission rate simulation runs isolate the impacts of the expected future reductions in fleet emissions while holding all other variables constant. Of course, the actual levels of auto emissions

Table 2

POLICY SIMULATIONS AND AREAS OF APPLICATION

Policy Simulation	Area of Application in Boston	Area of Application in Los Angeles
a. Reducing Mobile Source Emissions		
1. 1976 Auto Emmissions	all zones	all zones
2. 1980 Auto Emissions	all zones	all zones
b. Discouraging Automobile Trips to the Central Area		
3. Parking Surcharge	zones 1,2,3,6	zone 1
4. Local License	zones 1,2,3,6	zone 1
5. Prohibition	zones 1,2,3,6	zone 1
6. Prohibition-Congestion	1,2,3,6,4,5,9,17,20, 22	1,2,14,15,16, 31,45,46
c. Making Transit More Attractive		
7. Transit Fare Reduction	all zones	all zones
8. Central Bus Lanes	-----	1,2,14,15,16 31,45,46
9. Transit Performance Improvement	all zones	all zones
10. Rapid Rail Extension	9,23,33,20,22,21,38, 11,12,13,14,50,51,52	-----
d. Regulating Traffic Flow		
11. Faster Central Auto Speeds	zones 1,2,3,4,6,20	zone 1
12. Slower Central Auto Speeds	zones 1,2,3,4,6,20	zone 1
13. Faster Regional Auto Speeds	all zones	all zones
14. Slower Regional Auto Speeds	all zones	all zones
e. Promoting Car pooling		
15. Increased Auto Occupancy	all zones	all zones
f. Controlling Urban Development Patterns		
16. Centralize Activities	all zones	-----
17. Decentralize Activities	all zones	-----

in future years will also depend upon the nature of metropolitan growth which actually occurs.

These policy simulations both in Los Angeles and in Boston indicate that the predicted impact on air quality of reducing automotive emissions is dramatic. Aggregate mobile emission rates and pollutant concentrations of CO, HC, and NOX in the air quality control regions fall substantially. For the 1976 fleet aggregate mobile emissions fall by fifty-two percent for CO, by sixty-two percent for HC, and by twenty-two percent for NOX; for the 1980 fleet aggregate mobile emissions are down eighty-one percent for CO, eighty-five percent for HC, and sixty percent for NOX.

Carbon monoxide is emitted primarily by motor vehicles, so the forecast reduction in zonal CO concentrations closely parallels the forecast reduction in automotive emissions both in Los Angeles and in Boston. For CO the decline in concentrations is not exactly proportionate to the decline in aggregate emissions because the standards are projected to have a greater effect on running than on cold start emissions. Since emissions of hydrocarbons and oxides of nitrogen are produced by both stationary and mobile sources, the forecast reduction in concentrations for these pollutants shows considerably more variation.

Because of differences in the composition of stationary source emissions in the two cities, reducing mobile emissions of hydrocarbons generally produced a greater reduction in zonal concentrations of hydrocarbons in Boston than in Los Angeles. A comparison of NOX concentrations in the two cities revealed that the opposite is true for NOX emission reductions. Nevertheless, in Los Angeles, the zones with the

highest concentrations of stationary source hydrocarbons had their hydrocarbon concentrations reduced thirty-five percent by the 1980 fleet emission rates. In the Boston zones with high concentrations of NOX from stationary sources, a twenty-five percent reduction in NOX concentrations was achieved by the 1980 fleet emission rates.

b. Discouraging Automobile Trips to the Central Area

Three types of local vehicle restraint schemes were simulated: parking charges, local licenses, and automobile prohibition. The restraint areas were defined as zone 1 in Los Angeles and zones 1, 2, 3, and 6 in Boston. These zones constitute a generous definition of the central business districts (CBDs) in the two cities.

The parking charge simulation raised the average cost of parking by fifty percent. The area license scheme placed a 25¢ toll on all trips which originate outside and enter the restraint area. Trips which originate within the area are exempt from the local license toll. Outright prohibition is similar to licensing except entry cannot be purchased at any price.

Although traffic flow changes do not endogenously change speeds on the transport network in TASSIM, speed changes can be introduced exogenously if there is reason to believe such changes will actually occur. For these simulations it was assumed that the reduced traffic increased auto speeds by ten percent in the restraint area for all policies except the parking charge, which does not reduce through trips. In addition, a second prohibition simulation slowed auto speeds by 10 percent in zones adjacent to the restraint area to represent the congestion caused by traffic diverted from the restraint area.

The results of the simulations were generally similar for the two cities. On an aggregate level, all policies increased transit use and reduced auto vehicle miles traveled. All simulations except the Los Angeles prohibition-with-congestion run reduced aggregate mobile emissions. This exception occurred because the congested adjacent area modeled in Los Angeles was large, and lower speeds in that area caused an increase in CO large enough to offset the savings from the reduction in vehicle miles traveled.

The aggregate improvements from the restraint policies are not large, however; the really substantial concentration decreases occur only in the zones directly affected by the policies. The licensing and prohibition schemes induce auto tripmakers destined for the CBD to switch to transit outside the restraint area and to complete their trips on the public mode. In addition, they divert through trips into adjacent zones to transit, but they do not divert through trips from the restraint area.[8] As a result, while all three traffic restraint policies significantly improve air quality in the restraint area, licensing, and prohibition produce greater improvements than parking charges.

It should be noted, however, that the auto trips redirected to transit, the diverted through trips, and, if it should arise, the increased congestion in neighboring zones caused by traffic-restraint policies, all can have detrimental effects on the air quality in these adjoining zones. Through these effects were often not large, they are a reminder that central area restraints can have certain undesirable air quality impacts.

c. Making Transit Alternatives More Attractive

A variety of transit improvements are often proposed as techniques for improving air quality in central sections of metropolitan areas. It is argued that making the transit system more attractive will divert drivers from their vehicles, reduce vehicle miles travelled (VMT), decrease auto emissions, and improve air quality.

The TASSIM model has simulated in Boston and in Los Angeles the effects on air quality of transit fare reductions and performance increases. Fare reductions of ten percent in Boston and twenty percent in Los Angeles were tested. In addition, twenty-percent reductions in total time required for transit trips were examined. A central area bus lane scheme was tested in Los Angeles. Finally, the impact of extending the rapid rail system in Boston was considered.

The results conformed to results of other studies of transit ridership. Responses to fare reductions were less than responses to proportionate travel time reductions. Price elasticities for transit are low--about -.4 for Boston and -.2 for Los Angeles. In other words, a twenty-percent fare reduction increases transit ridership by about four percent in Los Angeles and eight percent in Boston. Transit level time elasticities are typically larger--about -1.0.[9]

Because the percent of trips on transit is not large, the fare reduction and performance improvement produced barely perceptible reductions in the total number of auto trips made in the two analysis areas. The rapid rail extension in Boston did increase the share of transit trips, but was unsuccessful at improving regional air quality. The Los Angeles bus lanes improved air quality in the CBD, but resulted in imperceptible changes for the rest of the region. Since the transit systems in both cities provide more complete service in the central

areas, fare reductions and improved transit travel times did generate some noticeable improvements in central area air quality. For example, in some zones concentrations of CO were reduced as much as thirty percent by the overall transit travel time reductions.

Although there is no persuasive evidence that aggregate emissions are reduced by transit enhancement policies, substantial air quality improvements can be generated in specified areas. Reductions in transit travel times have somewhat greater effects than fare decreases or extensions of transit service in producing these local improvements.

d. <u>Regulating Traffic Flow</u>

The emission rate of automobiles per vehicle mile is a function of vehicle speed: emissions of CO and HC decrease sharply with speed, whereas emissions of NOX increase slightly as vehicle speed increases. Thus, policies that increase vehicle speeds have often been proposed to improve air quality. Although increasing vehicle speeds could be envisioned as a region-wide policy, it is most frequently advocated for centrally located, highly congested zones where average speeds are quite low, ranging from ten to fifteen miles an hour. Increasing speeds from such low levels can significantly reduce emissions of CO and HC per vehicle mile. Conversely, policies that may reduce auto speeds could be expected to degrade air quality.

Four simulations that investigated alterations in vehicle speeds were carried out for both Boston and Los Angeles. The first policy simulation increased automobile network speeds by ten percent in a small central area, and the second policy simulation applied a speed

decrease in the same central area. The third and fourth policy simulations respectively increased and decreased automobile network speeds in the entire region.

The aggregate statistics from the region-wide speed changes are consistent with expectations. Faster auto speeds reduce total emissions of CO and HC and increase slightly those of NOX.[10] Slower regional auto speeds have the opposite effect. However, changes in the spatial pattern of pollutant concentrations in Boston differ from those in Los Angeles. When auto speeds fall in Boston, the transit system attracts enough tripmakers to reduce concentrations of all pollutants in the central zones. Conversely, faster auto speeds lure away enough transit riders to increase concentrations of all pollutants in the central zones. In Los Angeles, the transit system carries proportionately fewer riders. Therefore, very little shifting occurs, and changes in central area concentrations are similar to regional changes.

The small area speed changes reveal another difference between Boston and Los Angeles. In both cities, local auto speed improvements discourage transfers from automobiles to transit. Auto speed reductions have the reverse effect. However, central area speeds are substantially slower in Boston than in Los Angeles. Given the relationship between speeds and emissions, a ten percent speed change in Boston produces a smaller percentage change in individual vehicular emissions than a ten percent speed change in Los Angeles. Therefore, changes in through trips and switching behavior offset changes in individual vehicular emissions in Boston but not in Los Angeles. The small area speed change simulations suggest that in Boston speed improvement strategies that are not implemented with vehicle restraint systems that restrict auto access to the area will have a high probability of failure.

In Los Angeles, they will generate slight changes in the expected direction, i.e. speed increases will improve air quality. However, it is clear that in Los Angeles, as well as in Boston, adding vehicle restraints when auto speeds are increased would generate greater gains.

e. Increasing Automobile Occupancy

It is well known that the average auto vehicle occupancy rate is quite low relative to vehicle capacity. This knowledge generates enthusiasm for policies to promote more car pooling for work trips. However, because data on accepted and rejected car pooling opportunities are virtually nonexistent, few effective policies have been devised to promote higher vehicle occupancies. Although the TASSIM model yields few insights about the determinants of car pooling, it can be used to examine the air quality impacts of achieving specified higher levels of auto occupancy. When this policy is simulated by the model in the two cities, substantial and proportionate reductions in auto emissions and improvements in air quality are obtained.

Of course, car pooling can have effects on transportation demand that go beyond simply increasing the average number of persons per vehicle. If car pooling occurs mainly on work trips, some workers will leave their automobiles at home several days during the week. These automobiles will then be available to other household members who may use them for shopping, school, or personal business trips. This greater availability of automobiles caused by car pooling might increase the total number of trips made, thereby offsetting some of the gains from car pooling. A partial test of this phenomenon, using the Boston version of TASSIM, indicated about a third of the expected air quality improvements are eliminated if car pooling increases auto

availability and more trips are made. The number of trips may remain the same, however, if the primary effect of car pools is to reduce the number of second and third cars owned. If the increases in automobile occupancy are achieved primarily through car pooling for work trips in central zones, and if the additional trips generated are primarily suburban shopping trips, then significant central area air quality improvements may be generated which are greater than the reductions in aggregate emissions. These considerations suggest that increased car pooling will not reduce automobile emissions uniformly over an air quality control region.

f. <u>Controlling Urban Development Patterns</u>

Some analysts, recognizing the interrelationship among land use patterns, travel activity, and stationary emission source locations, have proposed that land use regulations and controls be used as policy variables to improve urban air quality. Implementing land use plans to improve air quality is difficult, however, because little systematic knowledge is available about the exact relations between land use and air quality. Therefore, there is little basis in terms of air quality effects for accepting or rejecting alternative land use plans.

The TASSIM model was used to investigate one particular hypothesis about the impact of land use on air quality in Boston. This hypothesis states that higher density development of central portions of metropolitan areas improves air quality by reducing the demand for transportation and particularly automobile travel and, conversely, that existing patterns of low density, dispersed development degrade metropolitan air quality by encouraging the use of autos and requiring longer trips.

Two scenarios of altered land use patterns were simulated with the Boston model. The first increases centralization in the air quality control region by moving twenty percent of the residences and employment located in the outlying ring to the central core. The second scenario increased the extent of decentralization by moving twenty percent of the core's population and employment to the ring.

Most of the transport statistics produced by these simulated land use changes conform to the expectations which suggested the policies. Centralization reduces the aggregate number of trips, trip lengths, and the aggregate levels of emissions of auto vehicles. Decentralization has the opposite effect on all of these variables.

The simulations predict certain other changes which have opposite air quality implications, however. Centralization causes average auto speed to decrease. In addition, auto trips to and through the downtown area increase. The net impact of these forecast changes from centralization is to increase concentrations of auto pollutants in the central zones by as much as 10 percent, and to decrease concentrations in the ring zones by as much as seven percent. Decentralization causes opposite, but more dramatic concentration changes, since the movement of jobs and people is absolutely greater.

Reducing low density peripheral development and increasing central densities may not be an effective policy for solving urban air quality problems unless it is combined with other policies that fundamentally change patterns of transportation usage. If altering patterns of travel usage is difficult, then the process of dispersal and decentralization underway in most metropolitan areas may actually help to reduce the concentrations of primary pollutants in the centers of those areas.

Since decentralization increases total emissions of HC and NOX, however, it might increase concentrations of secondary pollutants such as photochemical oxidants in some cities.

g. Summary of Policy Effectiveness

The preceding sections have provided a descriptive overview of the predicted effects that several control policies have on the transportation system and on air quality in two cities. This section summarizes the results of these policies with a region-wide and local "effectiveness' measure to quantify the impact of the policies on air quality.

The first measure of effectiveness, the regional effectiveness index described in the first section, measures region-wide violations of the ambient air quality standards in terms of person-times of exposure to pollutant concentrations that exceed the standards.[11] The second measure of effectiveness focuses on the local or target area improvement in air quality; it is the percent change in the concentration of carbon monoxide in the central business districts of Los Angeles and Boston. Many of the transportation control policies that were simulated were applied only to the two cities' central areas, so this latter measure indicates how much locally applied policies can improve local air quality. The regional effectiveness index, on the other hand, better summarizes the impact of policies on the entire air quality region.

The local regional effectiveness of the policies are compared in figure 4 for Boston and in figure 5 for Los Angeles. The origins of these figures represent the benchmark run. Not easily read from the

FIGURE 4

LOCAL AND REGIONAL EFFECTIVENESS OF POLICIES FOR BOSTON

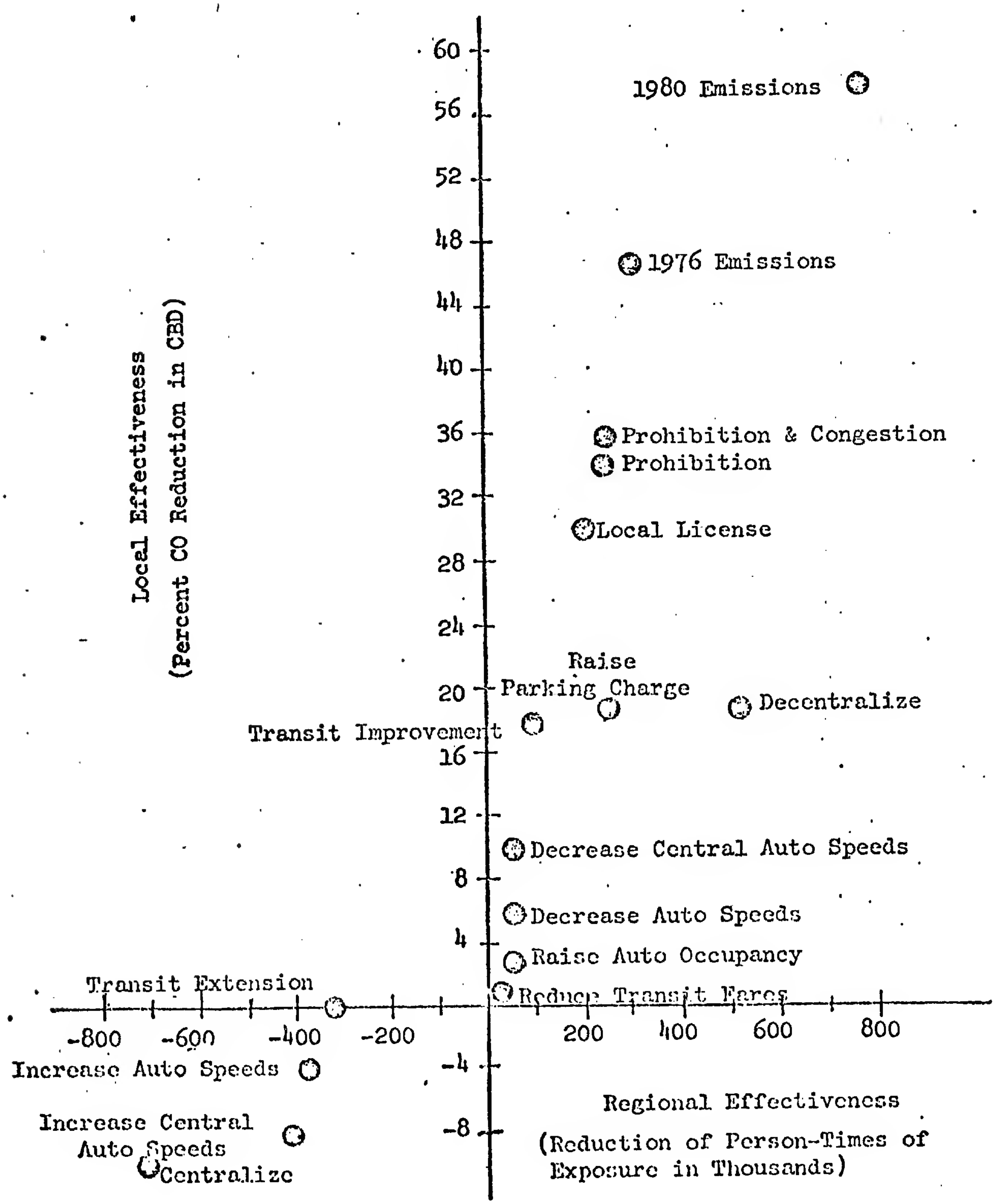
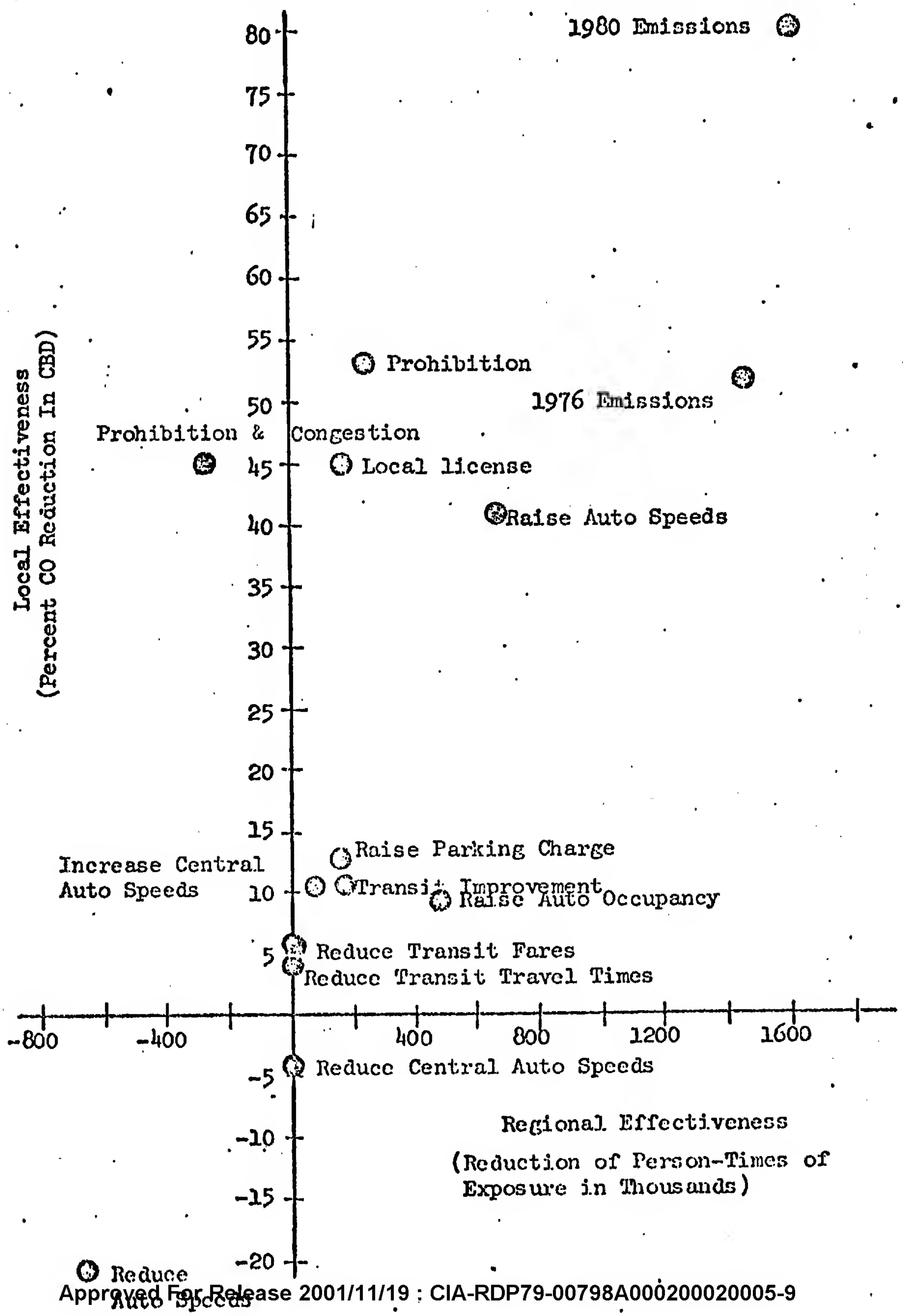


FIGURE 5

LOCAL AND REGIONAL EFFECTIVENESS OF POLICIES FOR LOS ANGELES



diagrams themselves is the observation that local rather than regionally directed policies have the greater relative impact, with the exception of the emission rate reductions. These diagrams highlight the differences between the policies in Los Angeles and Boston. Although many of the policies differ in their relative effectiveness, the only policies that act in opposite directions in the two cities are the speed increase and the speed decrease policies, as described under heading d. above. The only other policy that degrades air quality is the centralization policy simulated for Boston. Extending the transit system produces a negligible local improvement, but some regional air quality degradation in Boston. All other policies to discourage auto use and to increase transit ridership have positive regional and local effectiveness. As expected, policies applied only to the central business districts have greater local than regional effectiveness.

Finally, figures 4 and 5 illustrate how the effectiveness of emission reduction compares with the transportation control policies. The 1976 fleet emission factors are more effective locally and region-wide than any of the transportation policies simulated with 1970 fleet emission factors except for the regional improvement from decentralization simulated for Boston. The 1980 fleet emission factors produce substantial improvement in air quality relative to 1976. Because fleet emission factors will decline so markedly during the late seventies and early eighties, it is probable that vehicle control strategies will be required to meet air quality standards in most metropolitan areas for an interim period of only a few years. Of course, in the longer run, increases in automotive usage could mean that localized control strategies may have to be resurrected.

4. COST EFFECTIVENESS OF POLICIES IN BOSTON AND LOS ANGELES

Transportation control policies that have similar impacts may have widely different costs. It is necessary to determine costs for each policy to make judgments about their relative efficiency. Many simplifying assumptions must be made to evaluate each policy, but even crude estimates of the cost and effectiveness of various policies allow choices to be made among them. In this section resource costs for several of the policies simulated for Boston and Los Angeles are roughly estimated and compared with the two effectiveness measures used in the previous section.

Many components of the costs of the simulated policies can be derived from the model itself. For example, when transportation policies are applied, the model estimates changes in the use of automobiles, in transit ridership, and in the amount of time devoted to travel, which then have to be valued. The model provides no insights about the capital or operating costs required to implement a policy. These figures must be estimated separately, but estimating these cities costs for some policies is difficult because there are few examples of their use.[12]

For each policy the model forecasts total vehicle miles travelled (VMT) the number of round trips that originate on transit, the number of passenger miles travelled (PMT) on transit by persons who transfer from auto to transit for a portion of their trip, and the number of hours per weekday that people spend travelling. Table 3 displays these four quantities from the benchmark runs for Boston and Los Angeles. Since policy costs are derived from changes in each of these four measures, the absolute levels displayed in table 3 provide a base for determining absolute changes.

Table 3

BENCHMARK TRAVEL MEASURES

FOR BOSTON AND LOS ANGELES

	Boston	Los Angeles
VMT	21,960,336	165,936,000
Transit Originating Round Trips	348,252	190,052
Transit PMT From Auto	446,367	326,321
Hours of Travel	2,040,852	6,450,413

Source: TASSIM simulations

The changes relative to the benchmark run in each of the four transportation measures for each policy simulation are displayed in tables 4 and 5 for Boston and Los Angeles respectively. For example, the first entry in the VMT column in table 4 indicates that vehicle miles travelled relative to the benchmark declined by 71,000 in the parking-charge simulation for Boston. For the same policy relative to the benchmark run, transit originating trips increased by 9,008, transit passenger miles travelled (PMT) from auto originating trips increased by 379,167, and total person hours spent travelling rose by 8,835. The pattern of changes in the travel measures displayed in tables 4 and 5 is similar. If a policy increased a particular travel measure in Boston, it usually increased the travel measure in Los Angeles. The absolute changes in the travel measures are typically larger in Los Angeles than in Boston but, when scaled by the size differences of the two cities, they are generally similar. The most significant differences between the two cities involve the changes in transit originating trips. These differences are attributable to structural differences between the alternative modes of travel in the two cities: Boston has a much more highly developed transit system than Los Angeles.

The travel measure changes displayed in tables 4 and 5 were multiplied by costs or by prices to value them in dollar terms. VMT are valued at six cents a mile, an amount that approximates incremental operating costs for autos. Transit trips are valued at sixty cents each, which approximates the average cost per passenger trip in the Boston and the Los Angeles areas. Finally, hours of travel time have been valued at $1.75, which is one-third of the median wage rate in

Table 4

CHANGES IN TRAVEL MEASURES RELATIVE TO BENCHMARK

FOR SEVERAL POLICIES APPLIED TO BOSTON

	VMT	Transit Originating Trips (PMT)	Transit PMT from Auto	Hours of Travel
Parking Charge	-71,000	9,008	379,167	8,835
Local License	-220,000	-1,297	330,685	17,362
Prohibition	-265,000	592	375,968	21,177
Prohibition & Congestion	-444,000	62	403,721	15,795
Fare Reduction	-247,000	9,035	72	19,018
Improved Transit Performance	403,000	21,632	683,345	25,046
Transit Extension	22,000	40,905	44,203	17,513
Raise CBD Auto Speeds	166,000	1,309	-88,342	-3,247
Lower CBD Auto Speeds	-166,000	-1,276	105,348	1,003
Raise All Auto Speeds	659,000	-5,060	-96,110	-110,434
Lower All Auto Speeds	-613,000	4,307	134,601	106,211
Increased Auto Occupancy	-1,975,000	0	0	0
Centralize	-1,111,000	36,628	40,598	-13,420
Decentralize	1,320,000	-57,606	-88,330	-15,575

Source: TASSIM simulations

Table 5

CHANGES IN TRAVEL MEASURES RELATIVE TO BENCHMARK

FOR SEVERAL POLICIES APPLIED TO LOS ANGELES

	VMT	Transit Originating Trips (PMT)	Transit PMT from Auto	Hours of Travel
Parking Charge	-2,198,000	4,380	3,127,652	100,417
Local License	-1,595,000	436	3,017,695	60,070
Prohibition	-2,108,000	-864	3,901,766	72,350
Prohibition & Congestion	-2,080,000	2,679	4,118,656	201,292
Fare Reduction	-258,000	17,974	-1,541	3,106
Central Bus Lanes	-2,546,000	23,000	2,600,000	-2,332
Improved Transit Performance	-3,302,000	58,703	3,490,186	-16,007
Raise CBD Auto Speeds	260,000	-491	-326,231	-23,792
Lower CBD Auto Speeds	-211,000	210	364,386	16,808
Raise All Auto Speeds	468,000	-6,640	-326,231	-1,172,193
Lower All Auto Speeds	-478,000	3,876	509,599	629,898
Increased Auto Occupancy	-15,184,000	0	0	0

Source: [illegible]

the two cities.[13] Since the travel measures are for weekdays, the weekday valuations have been multiplied by 300 to transform them into annual values.[14] The results of these calculations are displayed in tables 6 and 7. For example, the first nonzero entry in the VMT column in table 6 is -$1,278,000. This is obtained by multiplying -71,000, the charge in VMT from parking charges in Boston, by $.06 and then by 300.

Tables 6 and 7 also display estimates of the capital and administrative costs of implementing most of the policies. The costs shown for emission reduction assumes that the average incremental cost per car (relative to 1970) of reducing emissions in Boston is $40 in 1976 and $65 in 1980, while in Los Angeles the figures are $50 in 1976 and $85 in 1980. Capital and administrative costs are not displayed for some of the policies because of uncertainties about policy implementation. For example, increasing the average occupancy of automobiles by ten percent above their current levels would reduce VMT proportionately, but estimating the cost of obtaining such an increase is difficult because the determinants of car pooling are not well understood. Similarly, it would be very difficult to estimate the costs of centralization for Boston or for Los Angeles.

The costs in tables 6 and 7 are plotted against the effectiveness of the various policies in figures 6 through 9. For most of the policies the local effectiveness diagram resembles the regional effectiveness diagram, and the Boston diagrams resemble the Los Angeles diagrams. For example, in both cities all effectiveness measures increase steadily from fare reductions to parking charges, to improvements in transit performance, to local licensing, and finally to prohibition.

Table 6.

ANNUAL DOLLAR RESOURCE COSTS FOR SEVERAL AIR QUALITY POLICIES

IN BOSTON (000's OF DOLLARS)

Policy	Capital Operating & Administrative	VMT	Transit Originating Transit	Auto Originating Transit	Travel Time	Total
1976 Emissions	43,000	0	0	0	0	43,000
1980 Emissions	71,000	0	0	0	0	71,000
Parking Charge	300	-1,278	-3,243	11,375	4,638	11,732
Local License	2,300	-3,960	-467	9,921	9,115	16,909
Prohibition	2,300	-4,770	213	11,280	11,118	20,141
Prohibition & Congestion	2,300	-7,992	22	12,112	8,292	14,734
Fare Reduction	0	-4,446	3,253	38	9,984	8,829
Improved Transit Performance	15,000	7,254	7,787	20,500	13,149	63,690
Transit Extension	70,000	396	14,726	1,326	9,194	95,642
Raise CBD Auto Speeds	1,000	2,988	471	-2,650	-1,705	104
Lower CBD Auto Speeds	0	-2,988	-460	3,160	527	239
Raise All Auto Speeds	?	11,862	-1,820	-2,883	-57,978	?
Lower All Auto Speeds	?	-11,034	1,550	4,038	55,761	?
Increased Auto Occupancy	?	-21,000	0	0	0	?
Centralize	?	-20,000	13,186	1,218	7,045	?
Decentralize	0	23,760	-20,738	-2,650	8,177	8,549

Table 7

ANNUAL DOLLAR RESOURCE COSTS FOR SEVERAL AIR QUALITY POLICIES

LOS ANGELES (000's OF DOLLARS)

Policy	Capital Operating & Administrative	VMT	Transit Originating Transit,	Auto Originating Transit	Travel Time	Total
1976 Emissions	197,000	0	0	0	0	197,000
1980 Emissions	334,000	0	0	0	0	334,000
Parking Charge	750	-39,564	1,577	80,389	20,781	63,933
Local License	2,750	-28,710	157	77,537	31,537	83,091
Prohibition	2,750	-37,944	-311	100,281	37,984	102,760
Prohibition and Congestion	2,750	-37,440	964	105,832	105,677	177,783
Fare Reduction	0	-4,644	6,471	-40	1,630	3,417
Central Bus Lanes	2,000	-45,828	8,280	64,820	-1,224	28,100
Improved Transit Performance	40,000	-59,436	21,133	89,693	-8,404	82,986
Raise CBD Auto Speeds	2,500	-4,680	-177	-8,378	-12,490	-13,865
Lower CBD Auto Speeds	0	-3,798	-76	9,354	8,824	14,304
Raise All Auto Speeds	?	8,424	-2,390	-8,378	-615,400	?
Lower All Auto Speeds	?	-8,604	1,395	13,107	330,750	?
Increased Auto Occupancy	?	-273,312	0	0	0	?

Source: calculated from TASSIM simulations

FIGURE 6

COST AND LOCAL EFFECTIVENESS FOR BOSTON

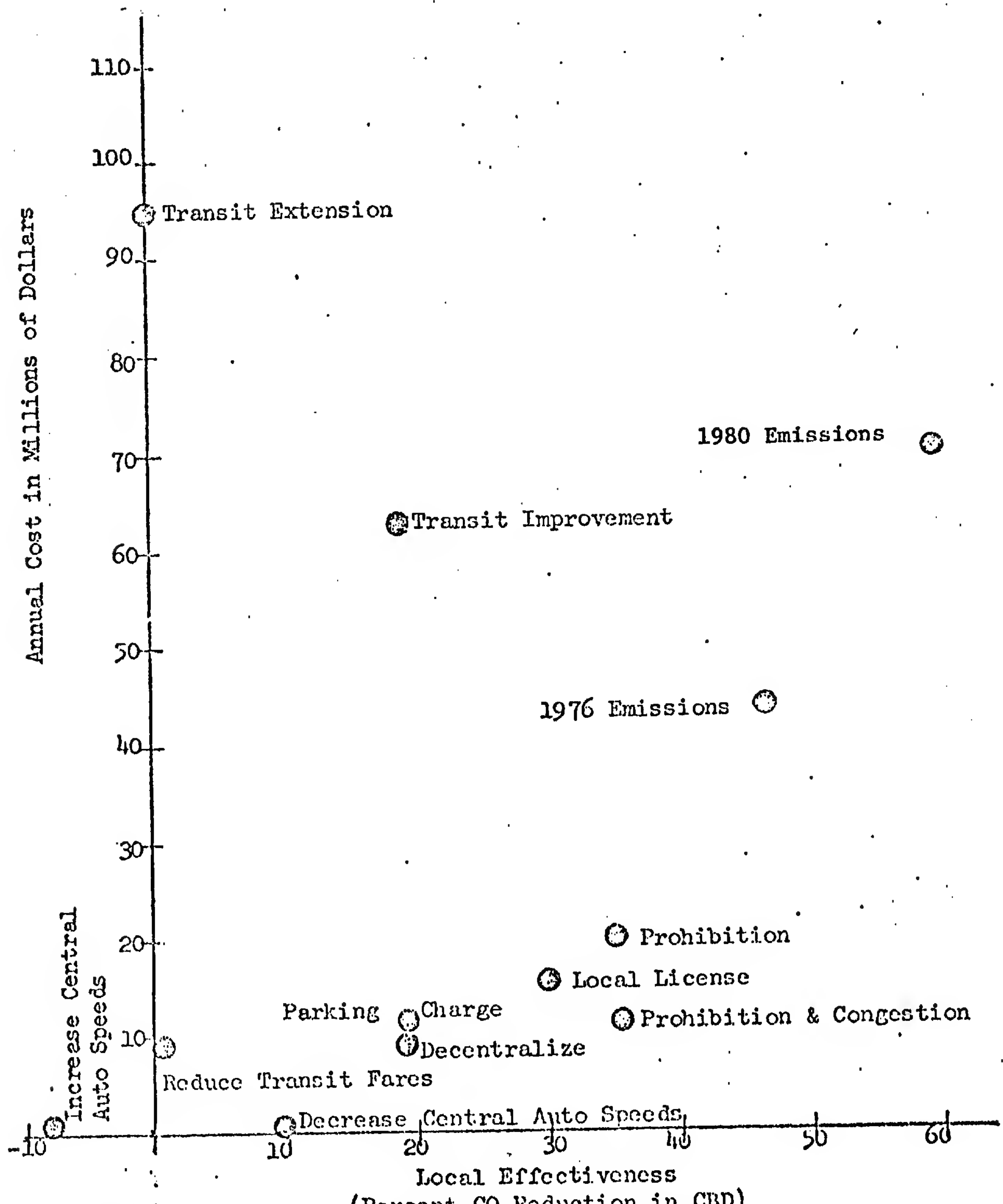


FIGURE 7

COST AND REGIONAL EFFECTIVENESS FOR BOSTON

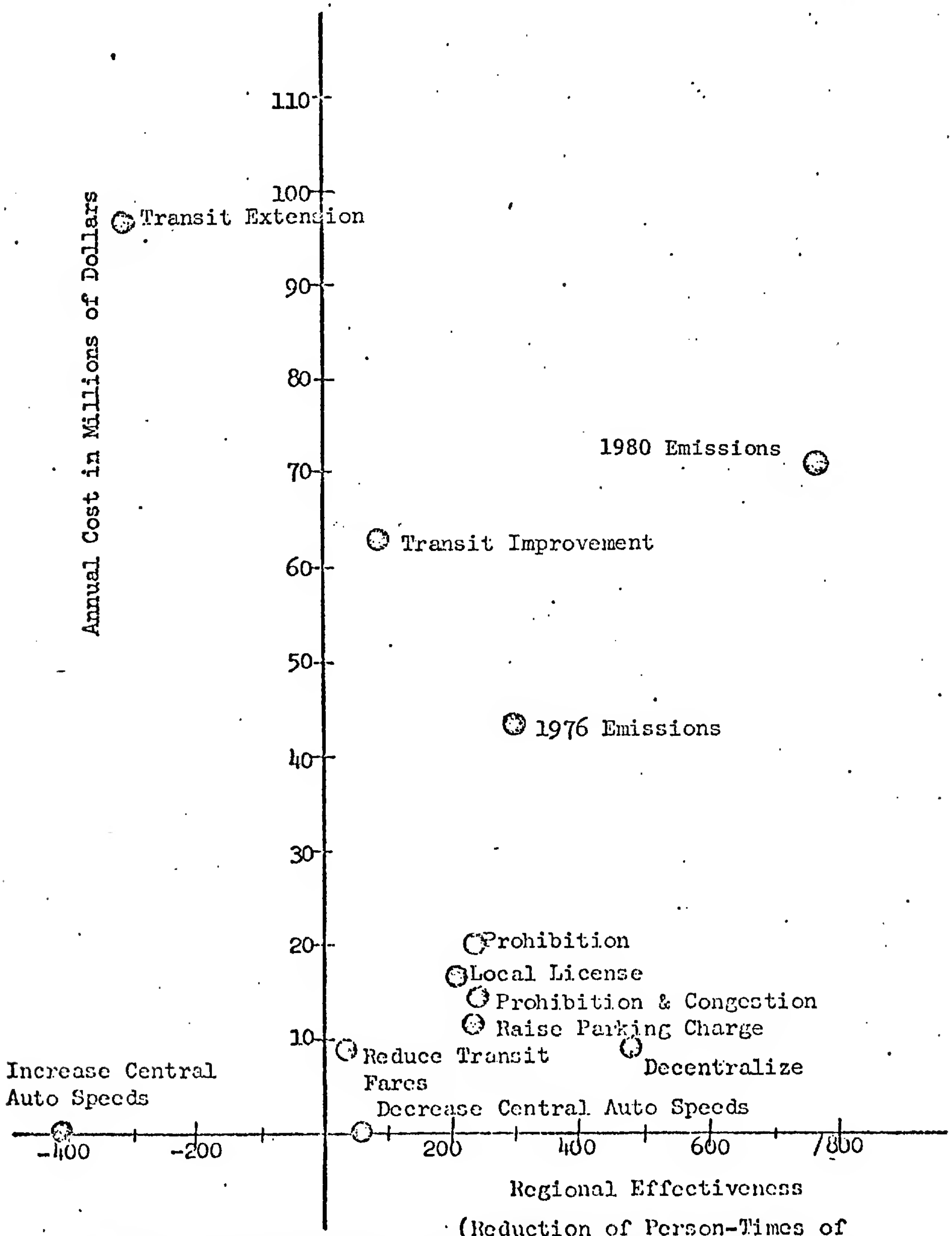


FIGURE 8

COST AND LOCAL EFFECTIVENESS FOR LOS ANGELES

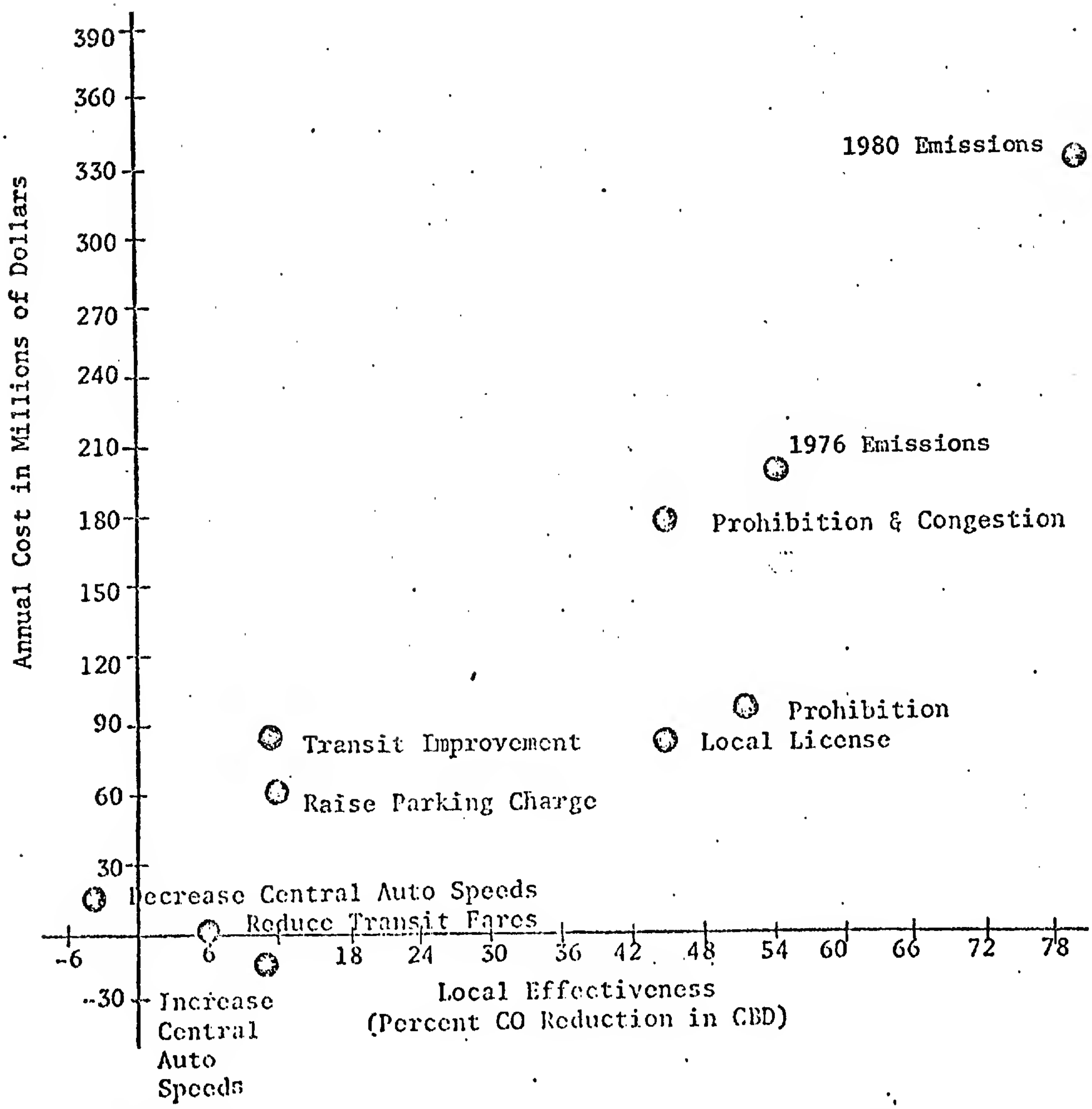
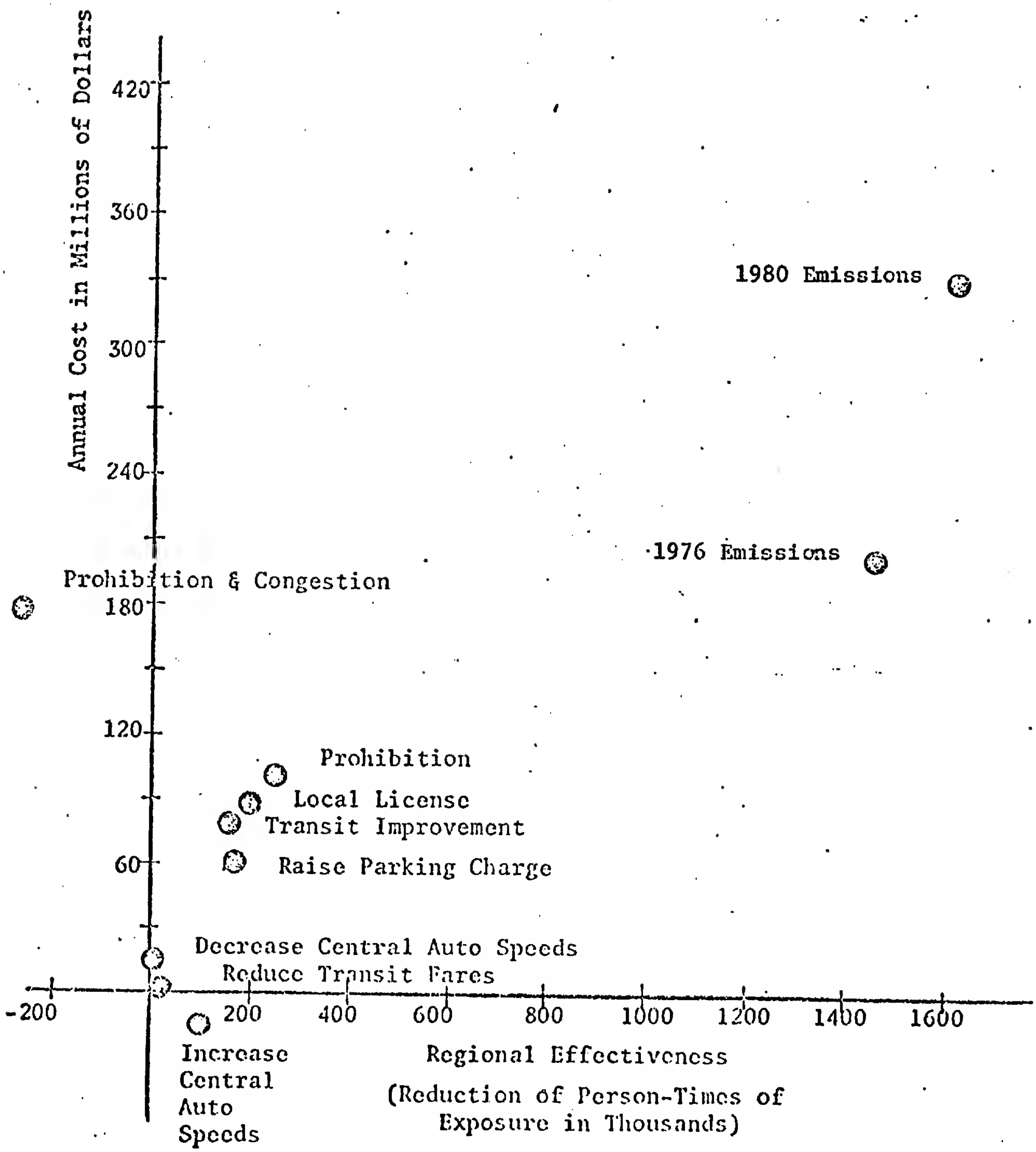


FIGURE 9

COST AND REGIONAL EFFECTIVENESS FOR LOS ANGELES



Prohibition and congestion are similar to prohibition in Boston, but are less effective than prohibition in Los Angeles. The major differences between Boston and Los Angeles occur in the central area speed changes, as explained earlier.

Reducing auto emissions to the legislated 1976 and 1980 levels both in Los Angeles and in Boston is cost effective relative to the transportation controls, although more clearly so in Los Angeles than in Boston. Several of the transportation controls buy small improvements in the effectiveness measures with expenditures that are also small, but large improvements in air quality are obtained only from the significant reduction in auto emissions over time.

The transit improvement policies other than reducing transit fares and implementing reserved bus lanes are not very cost effective in the two cities. In Los Angeles, similar improvements in effectiveness can be obtained at less cost from parking charges, whereas, in Boston, improving transit performance is dominated by local licensing or prohibition. The central bus lanes simulated in Los Angeles constitute one of the more cost effective transit policies. The transit extension simulated for Boston is not cost effective and actually decreases air quality as measured by the regional effectiveness index.

One final policy in the Boston simulation that deserves mention is decentralization. The implementation costs of this policy are assumed to be zero since the suburbanization of residences and employment will undoubtedly continue in the future. Decentralization performs very well on the measure of regional effectiveness because it reduces concentrations of primary pollutants. The model does not predict concentrations of photochemical oxidents, however, and

decentralization may increase levels of this pollutant by increasing aggregate emissions of NOX and HC.

5. CONCLUSION

This analysis of transportation control policies in Los Angeles and Boston has evaluated the air quality impacts of several policies by combining transportation and air quality forecasts produced by a computer simulation model with economic calculations of the costs of the policies. A comparison of the costs with the air quality impacts of the policies yields an assessment of the relative cost-effectiveness of the policies. This assessment can then be used as one means of choosing among alternative policies.

Although this paper has focused on air quality impacts, it is obvious that the transportation portion of the TASSIM model could be used to carry out preliminary analyses of other issues. For example, the effect on the transportation system of various land development policies could be approximated with the model: or preliminary estimates of ridership levels on extensions to transit systems could be provided. Of course, at the project level, more detailed analyses would be required, but the TASSIM model could help analysts identify those projects which merited further analysis. The model is inexpensive to operate and a large number of alternative policies can be compared fairly quickly at the regional level. Only those policies or programs which are particularly effective need be subject to project level evaluation. The use of TASSIM or other so-called "sketch planning" tools should help give analysts an overview of a wide range of policy alternative and help them to focus their project analyses more effectively.

APPENDIX: COMPUTATIONAL SUMMARY OF THE MODEL

The TASSIM model is typically executed on the computer as three separate computer programs named TASAQD, TASSIM, and TASMAP. The TASAQD program of the TASSIM model supplies data on air pollutant concentrations due to large point sources to the TASSIM program. TASAQD is a simplified version of the Martin-Tikvart Air Quality Display Model. It could be integrated into the same physical code as the TASSIM program. However, its isolation lowers the costs of TASSIM forecasts, since the majority of policy simulations look at strategies which have no effect on point source emissions. Therefore, the TASAQD program, under integrated operation, would simply reproduce its output over and over again. For Boston, the TASAQD program requires slightly less than two minutes of CPU (Central Processing Unit) time and about 2.3 minutes of total job run time per simulation of the 370/165. About 115 K-bytes of core storage are required to compile and execute the program at a cost of about twenty dollars on the Harvard-M.I.T. 370/165.

The TASSIM program performs virtually all of the operations described in the introductory discussion of the conceptual model. It requires .95 minutes of CPU time and about 2.7 minutes of total job run time per simulation on the 370/165. The Boston version of the TASSIM program requires about 156 K-bytes of core at a cost of about twenty-five dollars per run. Of course, these run times and storage requirements will vary if the dimensions of the model change. Faster run times could be obtained with H compilation and the use of object decks, and lower costs could be achieved by storing the input data on a user disk.

TASMAP is a supplementary program that graphically displays pollutant concentrations on maps for the entire study area and for the central core area. The maps are visual displays of information printed out elsewhere in the simulation run. The TASMAP routines could also be integrated into the TASSIM program, but using TASMAP to print out several maps at once is only slightly more expensive than printing out a single map. Therefore, it is efficient to complete several simulation runs and then to create maps for all of the runs at once. Thus, TASMAP, like TASAOD, is physically separate from the TASSIM program. Unlike TASAOD however, TASMAP is not essential to the TASSIM model, since it merely provides graphic displays of output produced by the TASSIM program. A typical map produced for Boston is displayed in figure 10. The Boston TASMAP program requires .9 minutes of CPU time, 1.4 minutes of total job run time, and 118 k-bytes core to produce twenty maps.

To adapt the TASSIM model to another city requires tables of travel times and trips, inventories of stationary emissions, meteorological parameters, vehicular emission rates, and socioeconomic characteristics of the population for calibration. Our experience with the Los Angeles version of TASSIM suggests that transferring the model to another city requires from six to eight man-months and about four hours of computer time on an IBM 370/165.

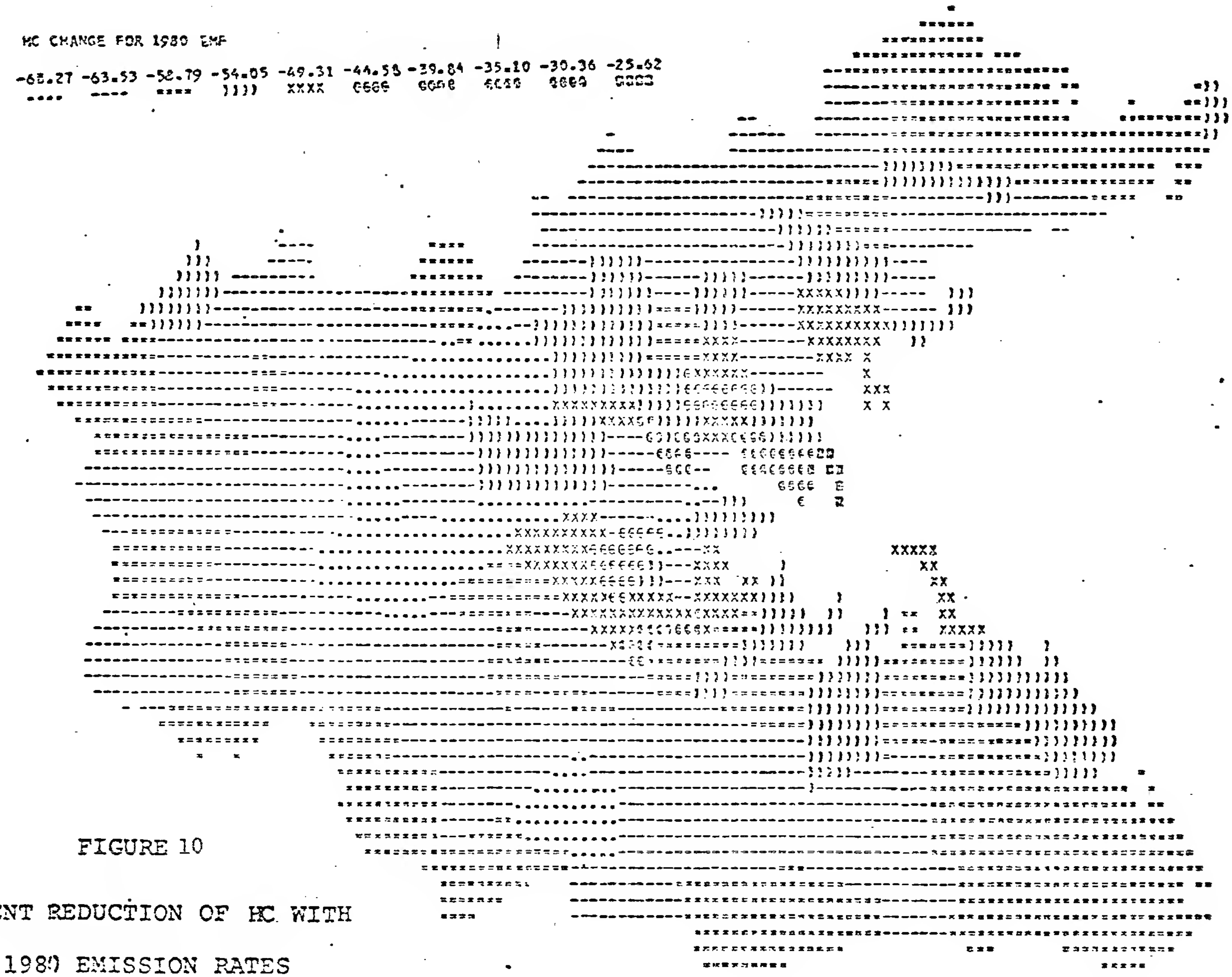


FIGURE 10

PERCENT REDUCTION OF HC WITH

1980 EMISSION RATES

NOTES

1. Extensive documentation of the TASSIM Model is available in Gregory K. Ingram, Gary R. Fauth, and Eugene Kroch, TASSIM: A Transportation and Air Shed Simulation Model, Vols. I and II, Final Report to U.S. Department of Transportation under contract DOT-OS-30099, May 1974.

2. These procedures are described in Urban Transportation Planning--Central Information, U.S. Federal Highway Administration, Washington, D.C., March 1972.

3. A "cold start" occurs when an engine that has cooled to the average air temperature is started. Cold start emissions differ significantly from the emissions of an engine running at its usual operating temperature.

4. The air source model is described in F. A. Gifford and S. R. Hanna, "Urban Air Pollution Modeling," presented at Second International Clean Air Congress, Washington, D.C., December 1970.

5. The point source model is described in D. O. Martin and J. A. Tikvart, "A General Atmospheric Diffusion Model for Estimating the Effects of Air Quality of One or More Sources," APCA paper No. 68-148, June 1968.

6. The diffusion models predict annual average concentrations of pollutants. A "Larsen transformation" is used to relate the predicted averages to the ambient standards. See Ralph J. Larsen, "A Mathematical Model for Relating Air Quality Measurements to Air Quality Standards," Washington, D. C., E.P.A. Report AP-89, November 1971.

NOTES (continued)

7. The emission reductions described reflect legislation in force as of January 1974.

8. The TASSIM model's price elasticity for parking is about -.3. Parking price elasticities typically range from -.3 to -.4. See Damian Kulash, "Parking Taxes for Congestion Relief: A Survey of Related Experience," Urban Institute Working Paper 1212-1, May 1973, processed.

9. These elasticities are similar to those estimated in other cities. See Gerald Kraft, "Free Transit Revisited," Public Policy, Winter, 1973.

10. Automotive emissions reductions from speed increases may be only temporary since an improved transportation system may attract more trips.

11. Recall that the simulations do not predict concentrations of photochemical oxidants.

12. Determining the costs of regional policies such as changes in regional auto speeds, increases in automobile occupancy rates, and alterations in land use patterns is beyond the scope of this study. For this reason we do not include them fully in our cost effectiveness analysis.

13. Empirical studies of travel behavior suggest that people value time at approximately one-third of their wage. See Michael E. Beesley, "The Value of Time Spent in Travelling: Some New Evidence," Economica, May 1965, pp. 174-195.

NOTES (concluded)

14. Weekday travel exceeds weekend travel, and the use of 300 rather than 365 provides crude compensation for this.

THE INTRODUCTION OF MATHEMATICAL-ECONOMIC METHODS AND COMPUTER TECHNOLOGY IN PLANNING AND MANAGING SOVIET TRANSPORTATION

B. S. Kozin *

The present period is characterized by an increasingly complicated overall transport system, rising intensity in the use of transportation plant and equipment, and closer connections between transportation and other branches of the economy. Successful functioning and development of the transportation system under these conditions depends to a large degree on the quality of management.

Planning is the foundation of management. Use of mathematical programming methods and computers have permitted the formulation and solution of a whole series of technical economic and planning problems with practical importance. This circumstance has assisted the appearance of theoretical research, including such problems as the formulation of a national economic plan considered as a type of extremal problem in mathematical programming. Theoretical research has shown the possibility in principle of constructing an optimal national economic plan. Under socialist conditions the possibility of working out an optimal plan becomes practicable. The following conditions assist this:

Presence of a unified institution influencing the planning activities of all branches and enterprises;

Noncontradictory goals for the functioning of different branches and groupings;

The possibility in principle of collecting and preserving all the necessary information on the status and functioning of the branches of the economy.

Introduction of economic-mathematical methods and computer technology in planning at the present time is directed toward:

---

*Director of Institute of Complex Transport Problems, Doctor of Technical Sciences

Obtaining optimal plan solutions for the basic problems of operating and developing the transportation system;

Achieving consistency among all decisions taken by plan organizations;

Carrying out by automatized means all the formal functions, most of which at the present time are carried out by planners (acquisition, compilation, and systematization of information, preparation of plan decisions, definition of their possible consequences, and so on). Automatization simply of the formal procedures met within the tasks of statistical and accounting record-keeping, material-technical supply, several planning tasks, and so forth, will, to some extent, permit a reduction in the management apparatus. Computers have already been used to solve this kind of problem in transportation for about twenty years. However a fundamental effect from using economic-mathematical methods and computers can be obtained in the productive sphere only through optimization of processes, more rational use of fixed and circulating capital assets. Besides, theoretical research and practical experience show that the effect of introducing economic mathematical methods and computers will be considerably greater if a system of tasks is solved, covering as a whole the defined functions of production, and not separate problems relating to certain partial aspects of these functions.

In the USSR three levels of automatized management of the transportation system are being worked out. On the first level Gosplan USSR is working out an automatized system for plan calculations in transportation (ASPRT) as one of the functional groupings for a subsystem in planning the national economy as a whole.

The basic tasks to be solved in the system ASPRT are:

Definition of the transport factor in the development of this branch of the economy (direct input coefficients in the interindustry matrix);

Definition of the final demands for transportation on the part of the population and the economy and the breakdown of the shipments among the various carriers;

Distribution of investment according to the direction of its utilization in order to attain a minimum outlay of national economic expenditure on the required shipments.

Each of these enumerated problems is in itself a large problem, their solution in a mutually linked manner will permit the working out of scientifically substantiated plans for the functioning and development of the transportation system of the country.

Each functional subsystem of ASPT, made up in accordance with general principles, and this includes the complex subsystem for transportation, is made up of several blocks. In the transportation subsystem, the following blocks are being worked out:

The general block

In the general block, problems are solved connected with defining the total demand for freight and passenger shipments and its subdivision among the various carriers.

The production block

In this block are worked out the integral indicators for the work of each means of transportation and, also, the volume of work assigned to each element for all means of transportation.

The science and technology block

In the science and technology block are worked out the possibilities for making use of the achievements of general scientific-technical progress

for the technical reequipment of transport; in this block are specified also the principal scientific problems and ways of financing their study.

The capital construction block

Transportation is one of the relatively capital-intensive branches of the economy. Transportation plant and equipment make up a significant part of the economy's fixed assets. Moreover, in the future it is intended that rather large capital investments will be assigned to the development of the transportation system. The effectiveness of new capital investment in transportation depends on how it is allocated. Solution of this problem is basic for the capital construction block.

The material technical supply block

In this block, the input requirements for rolling stock and equipment for each carrier are worked out, along with the necessary supplies of fuel (electric energy) and raw materials.

The block for labor and staff

In the labor and staff block we estimate the labor force required to operate the transportation system and, also, make up plans for training workers with the necessary qualifications.

The block for costs, profits, and profitability

In this block we calculate a summary measure of the activities of transportation as a branch of the system.

It is necessary to underline the fact that solutions for the problems of each block are linked together. Therefore, in order to obtain consistent answers, plan calculations cannot be carried through just once.

The plan solutions obtained through ASPRT for transportation by these methods both depend on, and have an influence on, the solutions that are reached in the other functional subsystems of ASPR. The mutual consistency of solutions is assured by the unified structure of subsystems, the general methodological foundation, and the unified technical means for processing information. The introduction of ASPR at all levels is connected with an improved methodological basis for solving plan problems through utilization of optimizing economic-mathematical models. In each of the blocks of ASPRT listed above, economic-mathematical models are worked out. Two of them--the planning of rail freight shipments and the planning of capital investments to increase rail line capacity will be described at the symposium.

The second level of automated systems for management, OASU (branch automated systems for management), is put together at the ministry level. At the present time, managing the work of transportation is handled by three union ministries (railroad, maritime, and aviation) and by republic organs of management for river and automobile transportation. Both in the union ministries and in republic GOSPLAN bodies, corresponding forms of ASU are being worked out. The planning block in OASU is one of the main ones. In this block the planning assignments worked out at a higher level are made more detailed. This disaggregation is carried out both with numerical indicators and through specifying "addresses" more precisely. Along with the tasks of transport planning, in OASU we solve in the first instance the problems connected with operating management of the shipment process, the planning of freight traffic, operational management of operating work, preparation of documents for performance standards, material-technical supplies, and various kinds of records. Each type of transportation OASU has its own specific characteristics reflecting the peculiarities in the functioning

and management of that kind of transportation. It must be noted that transport ASU differs from ASU for industrial branches of the economy. This is explained mainly by the fact that transportation is characterized by a more active role (compared with other branches) for the upper level in managing the productive processes. Such tasks as operational assignment of the freight car fleet throughout the railroad network, loadings at technical stations and along line sections, since they have first-class significance for railroad operations, can only be decided at the ministry level. Matters are the same with other means of transportation.

With the introduction of ASU for managing transportation carriers, the structural table of organization for management will also change and improve. Thus, while at the present time the railroads have a four-level system of management: ministry, railroad, division, line organization, with the introduction of ASU only three levels will remain: ministry-railroad-line enterprise. With the changed structure of management, there will be a corresponding redistribution of management functions. A three-level system of management will come into being on the other means of transportation also with introduction of ASU.

Within the limits of OASU, automated systems for reserving seats on passenger trains and aircraft are being created and developed. Special talks at the symposium will clarify these questions.

A third level of automated systems of management consists of ASU for technological processes and operations. Operating management of the work of freight classification yards, seaports, airports, etc., is connected with solving numerous and reciprocally interrelated problems in ASU of technical processes--this as a rule involves systems working in real time. Along with information processing equipment, a very important question for this kind

of ASU involves the collection and transmission of information. In connection with this, technology for collection (calculating equipment, preparation of documents, etc.) and for transmission (communication lines, commutators, etc.) must receive the necessary development. The main goal for operating planning consists of carrying out plans made up for a more extended period. In general, the principle for tying together all types of plans, beginning with operating plans straight through to long-term plans, consists of having the plans made at a low horizon of planning "concretize" the plans made at a higher horizon of planning.

OPTIMIZING MODELS FOR PLANNING THE OPERATION AND DEVELOPMENT OF A TRANSPORT NETWORK

I. T. Koslov *

The transport network is construed as a common graph whose elements are constituted by nodes and sections.

Each element of the transport network is characterized by a range of technical and operating parameters which define, with an acceptable degree of accuracy, its carrying capacity and the cost of moving cargo (passenger) within the boundaries of the element.

The performance of the transport network is determined by the load (quantity of cargo tonnage or number of passengers) on its component elements. Such a concept of the transport network is at variance with the notion of the transportation requirements of the national economy and the population. Determination of the transportation requirements of the national economy and the population is an independent problem in its own right, which lies beyond the scope of the present paper. But if one is to analyze the degree of utilization of the transport network carrying capacity, the loading of its principal elements comes in as a very handy tool. Thus, in the problem under consideration, planning of the transport network operation boils down to determining the loads on its component elements.

To outline the pattern of development of the transport network is to determine the kinds of reconstructive measures envisaged and the points in time when the network elements acquire additional capability.

Problem definition

The problem of determining the volume of operation and the ways of development of the transport network can be solved on the following premises:

*Division Chief, Institute of Complex Transport Problems and Doctor of Technical Sciences

1. The transportation requirements of the national economy that the transport network is to meet are expressed as follows:

   The pattern of relationships of the various kinds of cargo between the network nodes

   The number of passenger trains running over each network section

   The volume of the local cargo flows over the network sections

Each form of transportation requirements is liable to vary in time. In fact, the transportation requirement for the future is an indefinite quantity which only lends itself to probabilistic treatment. However, for the purposes of the present paper, it is assumed that this factor is a determinate quantity. Thus, the problem solution cannot but be treated as tentative. One can assume that the indeterminacy cannot detract from the usefulness of the problem to be considered.

2. The end goal of the transport system consists in meeting transportation requirements at the least cost. The transportation costs depend on numerous factors, some of which are relevant to the problem under discussion:

   The choice of cargo routes within the network

   The technical state of the main network elements

   The transportation pattern

3. Cargo may be carried from one node to another by quite a few different routes. Of course, the choice of the route for each individual batch of cargo never conduces to an optimal solution of the overall problem for the simple reason that the total transportation costs by and large depend on the loading of the elements encompassed by a given route. Furthermore, transport network elements with a limited carrying capacity often have to be excluded from the haulage routes for many cargo varieties.

The starting data for a model designed to determine the transportation costs consist of the loading of the network elements (nodes and sections). For this reason, it is the loading of the network elements, and not the cargo routes, that will form the unknown variables. Obviously, the set of the chosen routes uniquely defines the loading of the network elements.

4. The network elements are technically developed through reconstructive undertakings of many kinds. It is possible to limit the number of such undertakings, but only on the basis of special research. A research program of this type has been conducted, e.g., for railway network sections. The results of the study suggest that, of all the possible measures, only the following are practically significant:

   Extension of departure and reception station tracks as conducive to higher weight ratings of freight trains

   Substitution of automatic block signalling for the electric staff system or semiautomatic block signalling

   Construction of additional main tracks on open lines

   Substitution of electric locomotives for diesel locomotives

5. The type of traffic pattern significantly affects the economics and the degree of utilization of the available carrying capacity. In its turn, the traffic pattern by and large depends on the technological state of the network elements. However, some variables also exist; such as, for instance, the type of locomotives used. Since the locomotive stock is generally limited, it is impossible to use the most effective units over each section of the network. Thus, a problem arises as to the most advantageous distribution of locomotives among the elements of the transport network, which, naturally, cannot be solved unless the state of technology and the pattern of traffic are taken into account.

Mathematical definition of the problem

Let's introduce the main designations which will be used in the mathematical treatment of the problem.

Designations

s - number of the transport network sections (s = 1, 2, ..., S)

i - number of the transport network nodes (i = 1, 2, ..., I)

t - time (te/t_o, T_p/)

p - type of cargo (p = 1, 2, ..., P)

$R^+(i)$, $R^-(i)$ - totality of network sections going into and out of the node i

$x_{sp(t)}$ - flow of cargo of type p over the section s

$y_{i(t)}$ - operation of the node i

$\eta_{s(t)}$ - vector function determining the technological state of the section s

$\xi_{i(t)}$ - vector function determining the technological state of the node i

c(t) - traffic pattern

$A_{ijp}(t)$ - pattern of relationships reflecting the needs of the national economy for the transportation of cargo of type p

$N_s(t)$ - volume of passenger traffic over the section s

$F_s(t)$ - flow of local cargoes over the section s

K(t) - limit of investments into reconstruction

$N_n(t)$ - required carrying capacity of the transport network

$N_H(t)$ - actual carrying capacity of the transport network

$\pi(t)$ - passenger and freight transportation costs

A(t) - reconstruction costs

$\varphi(t)$ - weight function to account for the time value of expenditures.

The unknown variables in the problem considered are as follows:

section loads $x_{sp} = X$

node loads $y_i = Y$

section development $_s = M$

node development $_i = H$

traffic pattern c

The X and Y variables are determined by the choice of cargo routes. The section loads will be determined with a breakdown by cargo types conducive to more accurate technical and operating parameters of the operation.

In a general case, the node load should be broken down to differentiate between the components differing in the level of reclassification. Thus, for instance, for a railway node, not only the overall car flow must be determined, but also the transit car flow requiring no reclassification, the transit car flow which does require reclassification, and the local car flow. Such an elaborate characteristic of node loading can no longer be determined merely by cargo routes, for to characterize the transit flow one should break it down into the reclassified and nonreclassified flows which calls for the knowledge of the traffic pattern.

In the problem under consideration, the section and node loads are nonnegative functions of time.

The technical state of the sections and nodes may be characterized, with a sufficient degree of accuracy, by means of several parameters. As the latter vary with time, they indicate how the sections and nodes develop. The functions which define the development of the elements with time are characteristically discrete. Most elements have a limited number of states. Thus, for instance, the technical state of a railway section is characterized by hardly more than ten parameters, each of which can assume two or three values. Such a specific nature of the $M_s(t)$ u $V_i(t)$ function may be used to construct computational algorithms.

The traffic pattern varies widely depending on the particular kind of transport. Thus, the rail traffic system is by and large determined by the

distribution of marshalling operations among the stations and by the locomotive service.

The unknown variables are to satisfy three groups of constraints.

Group 1 includes constraints which reflect the compulsory nature of all transportation demands. Mathematically, these constraints may be represented as follows:

$$\sum_{s \in R_i^+} X_{sp}(t) - \sum_{s \in A_i^-} X_{sp}(t) = \sum_j Aijp(t) - \sum_j Ajip(t) \qquad (1)$$

$$i=1,2,\ldots,I;\ p = 1,2,\ldots,p.$$

Group 2 is composed of constraints which stipulate that the needed carrying capacity of the network should not exceed the actual value:

$$N^n(X,Y,N,F,t) < N^n(M,H,c,t) \qquad (2)$$

Group 3 comprises constraints whereby the needed reconstruction investments should not exceed the available level:

$$A(M,H,c,t) < K(t) \qquad (3)$$

The problem solution is evaluated on the basis of minimal reduced economic expenditures:

$$\int_{t_o}^{T} p\left[\mathcal{P}(XY,M,H,c,t) + A(M,H,c,t)\right]\varphi(t)dt = \min. \qquad (4)$$

Equations (1) to (4), together with the description of the function classes to which the unknown variables must belong, constitute the enunciation of the problem of traffic distribution and network development.

Problem analysis

The mathematical definition of the problem suggests that all the variables being optimized are functions of time. Assuming that time varies discretely, the unknown functions can be determined by finding the unknown values of the scalar variables.

By its type, the problem in question falls in the class of mathematical programming problems. It is impossible to define the problem more narrowly

with a view to using the well-developed techniques. Indeed, it is only constraints (1), which represent constraints of a linear programming problem in a network form, that have a straightforward structure, whereas the properties of the space assigned by constraints (2) do not lend themselves to clear-cut analysis. The left-hand and right-hand portions of equation (2) are calculated by special and fairly complex algorithms, which is also true for the left-hand portion of constraint (3).

All three groups of constraints (1) to (3) should be satisfied at each point in time independently. This requirement is used to construct a computational algorithm for solving the problem.

The properties of the functional to be minimized likewise elude analytical analysis.

However, one can assert that the functional to be minimized cannot be construed as a sum each term of which depends on just one moment of time. It is hence impossible to represent the general problem of traffic distribution and network development as a series of independent problems for a series of time sections.

Thus, the problem under discussion belongs to the category of mathematical programming problems of a general kind.

The actual Soviet transport networks have several thousand nodes and sections. Since the network carries several score aggregated types of cargo and the time base runs through at least three five-year periods, hundreds of thousands of variables will have to be determined if the problem is to be solved. Undoubtedly, an extremely complex computational task.

Problem solution

To obtain a general solution, the problem defined above will be broken down into the following parts:

Distribution of the required traffic about the network (R)

Determination of the required reconstructive measures for the network elements (V)

The solution is arrived at iteratively as follows:

$R_1 \to V_1 \to S_1 \to R_2 \to \ldots \to S_k$.

While solving the first part of the problem, both the technological level of all network elements and the traffic pattern are assumed to be known.

As the algorithm R is being realized, equalities (1) are satisfied for all time instants of the period in question.

The second part of the problem is concerned with the measures aimed at raising the technological level of the network elements. The traffic distribution and the traffic pattern are considered to be known, and the algorithm must be so constructed as to provide for satisfying inequalities (2) and (3) given in the mathematical definition of the problem.

The third part of the problem deals with the traffic pattern, the traffic distribution, and the technological level of the network elements being considered as invariable.

No proof has been offered to the effect that iteration conduces to an optimal solution. Moreover, the large dimension of the problem does not allow for too many iterations. It follows that an approximate solution is the only one that can be obtained by the proposed method.

The traffic problem algorithm can be employed as the algorithm R for determining the traffic distribution about the network. In fact, constraints (1) are exact analogs of the traffic problem constraints. The national economic expenditures on traffic, which are part of the general criterion (4), may be used as the yardstick for determining if the solution is indeed an optimal one. If the methods used to solve the traffic problem are to be

applicable here, the traffic expenditures must be represented as a sum of expenditures related to the network elements, and this kind of representation involves a certain error.

Numerous authors have demonstrated that the hypothesis which stipulates that expenditures grow linearly with the loading of the network element, is too broad, which calls for the use of more complex functional relationships between expenditures and loading. Analysis indicates that the real functions of expenditures all show an important property; viz., the curves are convex downward.

This made it possible to work out a special algorithm for determining the network distribution of traffic with nonlinear characteristics. It is a final algorithm conducive to a truly optimal solution.

The traffic distribution problem as stated must be solved many times for fixed moments of time within the period under consideration, so that the network element loads are actually obtained as functions of time.

The optimal development of the technology of the network elements is determined on the basis of the element load values established at the previous stage. Mathematically, this part of the problem could be reduced to the task of finding a minimum of the function of many variables which represent the sum total of all expenditures over the period under consideration. Such representation of the problem is made possible by the specific functions of development of the elements as indicated in the mathematical definition of the problem. When determining the optimal pattern of development, inequality (2) is taken into account for each element separately, whereas constraint (3) must be satisfied for all elements lumped together. To satisy this latter constraint, a heuristic algorithm is suggested which sets the order in which reconstructive measures will be taken, it being understood that the order

thus set takes effect only if limited investments are available for the development of the carrying capacity.

The optimal traffic pattern is determined by the results of the two previous solutions. Central to any traffic system (at any rate for the railway network) is the deployment of locomotives. Analysis indicates that this problem may be solved, given some assumptions, by linear programming techniques.

Special methods exist for scheduling and distributing marshalling jobs, but in this case approximate methods based on statistical data seem to suffice.

It is thus clear that the general problem of traffic distribution and technological development is actually a set of interrelated, highly complex, and computationally labor-consuming problems. Hence, the problem can only be solved through the use of efficient computer centers.

METHODS OF FIVE-YEAR PLANNING OF TRANSPORT-ECONOMIC CONNECTIONS: THEORETICAL DEVELOPMENTS AND EXPERIENCE WITH PRACTICAL APPLICATIONS

N. I. Mokrousova and Z. I. Mozgrina *

The volume of future traffic is crucial for many aspects of five-year planning. For this reason one of the first problems to be solved is one of planning the transport-economic connections, which furnishes the initial block of data on the future volumes and directions of cargo traffic.

From the economic point of view, the problem in question boils down to meeting the country's freight traffic demand at minimal cost within a preset production framework. In order to accomplish this task, optimization models are employed to plan the traffic volumes of the future.

Analysis indicates that for the purpose of long-term planning of transport-economic connections, all kinds of cargo may be classified in three broad groups:

1. Cargoes which are planned on the basis of territorial production-consumption balances, with the transport-economic connections of the suppliers and the consumers liable to change
2. Cargoes which are likewise planned on the basis of territorial production-consumption balances, but with the transport-economic connections remaining invariable
3. Cargoes which are planned without recourse to territorial production-consumption balances

The first-group cargoes may be broken down into two subgroups:

Cargoes assumed to be homogeneous (most types of timber, cereal, and petroleum products), which are singled out from the rough lists of bulk cargoes by virtue of their homogeneous processability as

* Senior Research Staff, Institute of Complex Transport Problems

well as consumer and other properties. For these cargoes, the transport-economic connections are determined by optimizing supplies by means of linear programming traffic problem for preset points of cargo departure and destination and for predetermined cargo volumes. The volumes of departing cargo are determined by balance calculations for specific products using the oblast (roughly equivalent to region) as a base territorial unit. For the purpose of this paper, this kind of model will be referred to as a one-product model;

Nonhomogeneous, but interchangeable cargoes, such as, for instance, fuel coal grades varying in calorific value or chemical fertilizers containing different levels of nutrients. These cargoes are likewise covered by oblastwise and stationwise departure and arrival schedules. The optimazation problems involved in the planning of supplies of interchangeable products are solved by the allocation problem method (λ-problem). This model will be called a λ-problem.

The second group includes mining and metallurgical products, construction materials, petroleum, coking coal, grinding products, and many other kinds of like cargo. Given certain established transport links, their quantitative characteristics for the long term are determined by traditional methods, and the traffic plan is evolved using an optimization model whereby cargoes (relationships) are superimposed on the transport network following the tree of the shortest (cheapest) routes. The model used for this category of cargoes will be termed a rigid-link model.

The third group includes several hundred agricultural products, foodstuffs, consumer goods, some chemicals, products of the woodworking and paper-and-pulp industries, metal items, machines, equipment, transport vehicles, industrial materials, and structural elements and the like.

The items coming under this heading are universally produced and consumed, so that it is impossible at present to compile territorial production-consumption balances for these lines of goods. Hence, special methods have to be devised for their planning.

Let us first dwell in some detail on the planning models of transport-economic links for the first two groups. The one-problem model, which is built around a linear programming traffic problem in network terms, is defined as follows.

Given is a transport network containing n vertices and N arcs, the arcs being designated as S = 1, 2, ..., N; the vertices i_s = 1, 2, ..., n and j_s = 1, 2, ..., n (i_s abd h_s are respectively the outgoing and the incoming vertices of the arc S). The cost of transporting a unit of cargo along the arc C_s is known.

Given are points of production and consumption of a homogeneous cargo. The production volume a_k is more than zero, the consumption volume a_k is less than zero, and a_k is equal to zero at the rest of the network points, or vertices, where the cargo is neither produced nor consumed (k = 1, 2, ..., n).

It is required to devise a plan of assigning suppliers to consumers and a corresponding traffic plan (x_s, S = 1, 2, ..., N) such that would involve a minimum of transportation costs $\sum_{S=1}^{N} c_s x_s \to min$

while satisfying the following conditions:

production-consumption balance

$$\sum_k a_k = 0, \quad k = 1, 2, \ldots, n; \tag{1}$$

flow equilibrium at the network vertices

$$\sum_{is = k} x_s - \sum_{js = k} x_s = a_k, \quad k = 1, 2, \ldots, n; \tag{2}$$

nonnegativity of traffic

$$x_s \geq 0, \; S = 1, 2, \ldots, N. \tag{3}$$

The information input for this model is constituted by an aggregated conventional transport network. At this stage, the railway network contains 550 nodes (vertices) and 850 segments (sections). Aggregation and substantiation of a conventional network are dealt with in numerous special investigations.

Reduced transportation costs and distances in kilometers were used as estimates of the network sections. The reduced transportation costs include the costs of operating trains, servicing them en route and maintaining installations, as well as the investments into the rolling stock and technological development. The costs and investments are incorporated in the part of the model which represents the traffic volume. The sectionwise indicators were differentiated for bulk cargo and oil products. Bulk cargo was defined on the assumption that all network sections were equally loaded in all directions and that the rolling stock was utilized to an average degree (80% utilization of gondola capacity and 90% utilization of boxcar capacity). For oil products, the assumptions were that the cargo was transported over all network sections and the tank cars were utilized to an average degree (80% utilization of tank car capacity).

While planning specific cargoes, the differences in the utilization of the rolling stock capacity are allowed for through the use of correction coefficients.

The estimate of each arc must incorporate the correction coefficient. In the course of computerized calculations, it was found advisable to use the main starting network variants for bulk cargoes and oil products, with the corrections being incorporated in the functional.

Experience with experimental and plan problems indicates that the network form of the traffic problem is the most convenient approach for planning practice. Its advantage over the matrix format consists in that it permits obtaining all elements of the traffic plan directly on the transport network (traffic over the network sections, relationships, and routes), evaluating the plan in terms of several criteria, and dispensing with the step of preliminary computation of the shortest and cheapest routes between specific suppliers and consumers, since problems for various kinds of cargo with any set of suppliers and consumers can be solved directly on the network.

The network problem makes use of two kinds of initial information, viz. constant and variable. The former includes data on the transport network which are prepared in advance and stored in the external memory of the computer. These data are used throughout the entire computation cycle, being refined and corrected from one cycle to another.

The variable information, incomparably smaller in volume, includes data on the cargo itself; i.e., on the arrivals and departures of the cargo at the vertices of the transport network.

For the one-product problem, there is a set of programs which, along with the transport problem program, also includes a program for planning the assignment of suppliers to specific consumers, the routes of cargo movement being indicated for each supplier-consumer relationship. The set further contains a program for determining the volume of work done in ton-kilometers. The programs provide for producing intermediate results (punched-card arrays) for use in crossfoot models.

The next model is one of planning of nonhomogeneous interchangeable cargo traffic. Here it is possible to use a linear programming allocation problem (λ-problem) which is mathematically defined as follows:

Given are the points of production i and consumption j ($i = 1, 2, \ldots, n$; $J = 1, 2, \ldots, m$), as well as the volumes of resources a_i and consumption b_j of interchangeable products. Also known are coefficients λ_{ij} which relate the value of production a_i to that of consumption b_j.

It is required to provide a product supply plan conducive to minimized transportation costs

$$\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} c_{ij} \to \min$$

on condition that

$$\sum_{j=1}^{m} x_{ij} < a_i, \; i = 1, 2, \ldots, n;$$

$$\sum_{i=1}^{n} x_{ij} \lambda_{ij} = b_j, \; j = 1, 2, \ldots, m;$$

$$x_{ij} > 0, \; \lambda\, i_j > 0,$$

where x_{ij} represents the supply of the products manufactured by the i-th supplier to the j-th consumer; and c_{ij} represents the cost of transporting a unit of products from the i-th supplier to the j-th consumer.

The interchangeable cargoes considered in planning practice allow using a somewhat more complicated traffic problem rather than special allocation problem algorithms. This is made possible by the fact that the coefficient λ for one of several suppliers has a constant value. In such a case, the coefficient λ_i is constant in each line of the matrix, if the traffic problem is solved by the matrix method. Thus, it is quite sufficient to divide the coefficients of the matrix c_{ij} by λ_i and solve a common traffic problem in conventional product units. For obtaining network results, the relationships thus found in natural units must be superimposed on the transport network.

For the network case, this problem can be solved in two stages. At the first stage, the transport problem will be considered for a multilayer network for the entire volume of products in conventional units, whereas the second

stage will deal with ordinary one-product problems for each individual product in natural units, the one-product problems being defined by the results of the general optimal solution. The network problem proved amenable to solution because a small number of heterogeneous groups of interchangeable products was considered: three kinds of fuel coal and two varieties of nitrogenous and phosphoric fertilizers. The number of identical layers for a multilayer network corresponds to the number of heterogeneous groups of cargo being considered. The estimates of the identical layerwise sections are set to be inversely proportional to the coefficient λ_i. A multilayer network with sequentially numbered layers can be produced automatically by means of a special program. The starting economic data for heterogeneous interchangeable cargoes are constituted by the departing and arriving cargo volumes in conventional and natural units at the vertices of the transport network. The set of programs of this model includes all programs of the one-product model and an additional accessory program of transformation of the starting network into a multilayer one by the preset values of λ_i.

The rigid-link model essentially contains a problem of distribution of predetermined cargo loads about the transport network. It is defined as follows:

Given is a transport network containing n vertices and N arcs, the arcs being designated by S ($S = 1, 2, \ldots, N$) and the vertices by i and j ($i, j = 1, 2, \ldots, n$).

The transportation costs over each are C_s are known. The traffic pattern is formed as a network of internodal relationships a_{ij}.

It is required to construct a traffic plan for the transport network (x_s, $S = 1, 2, \ldots, N$) providing for minimized transportation costs:

$$\sum_{s=1}^{N} C_s x_s \to \min.$$

This problem is built around a minimum path-length algorithm.

The cargoes planned with the use of this model are characterized by the need to retain, fully or partially, the transport-economic links evolved over a long time, while permitting certain changes in the distribution of production units and in the territorial production-consumption balances for the long term. Such conditions imposed on some links stem from technological requirements, specialization of some production units in terms of special raw materials or mixtures thereof, and from some other factors of a similar nature. While readily incorporated in traditional plans, they sometimes prove unmanageable for economic-mathematical models. To ensure automatic planning of transportation of this group of cargoes, the simplest possible solution was found; viz., to combine traditional methods of computation with automatic operations. Under this system, the interoblast and then internodal relationships accounting for rigid transport-economic links are handled manually by traditional methods, but the most labor-consuming computation tasks involved in the superposition of these relationships on the transport network and determination of the principal indicators of the plan are done automatically.

The set of programs in this model includes programs for planning the traffic along the preset pattern of relationships between the nodes, as well as calculating the work in ton-kilometers.

Summation programs were developed to produce the final results of the traffic operations, whereby resulting volumes can be obtained for any set of cargoes. Summation is carried out on the basis of the intermediate information furnished by the optimization models. The set of summation programs provides information on the overall traffic density over the transport network sections, the general and structural tables of internodal relationships, and

similar tables for interoblast, interrayon (rayon: a territory far larger than an oblast) and interrepublican cargo exchanges.

Experience with experimental and plan calculations suggests that the long-term traffic volumes are generally understated, both in terms of integral indicators and for the sectionwise flow density figures. The understatement was tentatively estimated at 1 to 10% for various kinds of cargo. The reasons for such understatement should be sought in the following factors:

Overaggregated long-term economic information worked out for an enlarged cargo nomenclature and large economic regions

An aggregated transport network including 550 stations instead of the 8,000 actual stations and covering a smaller area than in reality (113,000 km instead of the actual 135,000 km of overall route length due to the skipping of all dead-end sidings, underloaded sections and intranodal tracks)

The optimization models themselves which, under conditions of the above-described aggregation, provide ideal optimal plans free from all kinds, even rational, of cross-hauls of actually nonhomogeneous cargoes (in the one-product model)

In order to allow for the possible understatement of the integral and sectional traffic plan indicators, use can be made of corrections whose magnitude can be determined by comparing the actual data with the calculations and introducing an allowance for the long term.

All the above-described optimization models for determining transport-economic links are being currently introduced into the medium-term traffic planning practice.

The transport-economic links of the third group are planned on the basis of available statistics. The problem may be defined in the following manner.

Given are an interoblast and interrayon tables of relationships for a certain base year derived from reports, as well as an interrayon table of relationships and total freight turnover figure for a given group of cargoes for the period being now considered, the latter two parameters being found by a certain method. It is required to determine an interoblast table of relationships amounts to the corresponding interrayon relationships, and the freight turnover for a given group of cargoes is equal to the assumed given value Γ.

Mathematically, the problem is defined as follows:

Given: a matrix A each element of which (a_{ij}) characterizes cargo transportation from the i-th to the j-th oblast in the base year ($i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, n$).

a matrix L whose elements l_{ij} characterize the distances between the oblasts;

a matrix C each element of which (C_{rs}) represents the traffic from the r-th to the s-th rayon in the base year, and

$$C_{rs} = \sum_{i \in J_r} \sum_{j \in J_s} a_{ij},$$

where J_r is the set of all i's which belong to the S-th rayon ($r = 1, 2, \ldots, Z$; $S = 1, 2, \ldots, Z$);

a matrix D each element of which (d_{rs}) represents the traffic from the r-th to the S-th rayon in the year being planned.

A matrix b is to be found such that

$$\max_{i,j} \frac{b_{ij}}{a_{ij}\frac{drs}{Crs}} \to \min$$

on condition that

$$b_{ij} \quad a_{ij}, \quad V_{ij};$$

$$\sum_{i \in \tau_r} \sum_{j \in \tau_s} b_{ij} = drs, \quad V, S$$

$$\sum_i \sum_j b_{ij} \, l_{ij} = \Gamma,$$

where b_{ij} is the traffic from the i-th to the j-th oblast in the year being planned.

The above-defined problem is solved in the following way.

First of all, it is necessary to determine coefficients characterizing the increase in the volume of relationships between the rayons in the year being planned as against the base year. The coefficients thus obtained are used to determine a proportional table of interoblast relationships for the planned period, simultaneously checking whether or not the condition that a predetermined freight turnover is satisfied. If this condition is satisfied, the problem is all but solved. If not, then, maintaining the same general picture of departing and arriving cargo volumes, the relationships between the oblasts must be so changed as to arrive at the predetermined turover figure, taking pains to keep deviation from the proportional pattern of relationships within a minimum, because of the specific nature of this group of cargoes. With a constant departing volume, the turnover can be changed by varying the average length of haul. To this end, the relationships must be changed in the following way.

An additional table of relationships is found which actually amounts to the difference between the proportion and base tables. As the significance of relationships within each economic rayon diminishes, they are adjusted and finally broken down into pairs. The possible variation of the ton-kilometer index is assessed for each pair of adding the longest or the shortest distance depending on whether the turnover as a whole increases or diminishes.

The variation in the ton—kilometer index is calculated for each pair by multiplying the cargo quantity being transported by the distance differential. Integrating these variations for all pairs of cargo relationships for all economic rayons, we will determine the maximum value by which the turnover within a given table of relationships may be changed (increased or decreased).

In order to achieve the predetermined turnover value, it remains to determine the share of the possible maximum variation. This value is equal to the proportion of turnover deviation of the maximum value by which the turnover within a given table of relationships can be varied. The additional and then the starting tables of relationships may be varied in proportion to this share.

The cargo traffic density over the sections of the network under consideration cannot be determined merely on the basis of interoblast links. Investigation of the interoblast exchange table of relationships for the group of "other" cargoes for the year 1970 indicates that the intraoblast traffic accounts for thirty seven percent of all shipped goods, but in terms of turnover this kind of traffic is insignificant due to the short distances involved.

If the traffic density over the network sections is to be determined, a more detailed table of relationships is required, for, alongside oblast cities, there exist other big consumers (or suppliers) of "other" cargoes.

Analysis of the points of production and consumption of the group of cargoes under consideration allowed for determining several aggregated nodes for each oblast, finding their relative shares of the total turnover and, thereby, breaking up the relationships of the interoblast table into smaller units.

Knowing the production and consumption points and using the minimum path-length algorithm, the table of relationships can be superimposed on the

network being considered to find the traffic density for the "other" cargoes by the sections of the network.

The above-mentioned algorithm was used to perform experimental calculations on the basis of the 1970 railway reports. Analysis suggests that the calculations are entirely satisfactory for practical purposes.

PROBLEMS OF OPTIMAL PLANNING AND MANAGEMENT OF AUTOBUS TRANSPORT IN THE GEORGIAN SSR

G.G. Tsomaja *

1. Goal Description

The present level of motor transport technology, the ever increasing traffic volumes and the increasingly sophisticated technology and patterns of the transportation process, all emphasize the urgent need to improve the methods of planning and controlling the operation of the transport systems. The motor transport of the Georgian Republic is a highly complex system which will be progressively more and more difficult to plan and control unless modern control methods and technology are used. To meet this challenge, the "Avtotranstekhnika" /Automotive Transport Technology/ Agency of the Ministry of Motor Transport of the Georgian SSR has enlisted the cooperation of the Laboratory for Problems in Automatic Control and Computer Technology of the V. I. Lenin Georgian Polytechnic Institute to develop an automatic system for planning and controlling the operation of the passenger motor transport service. The program has set the following objectives:

Developing an effective network of routes

Selecting optimal types of the rolling stock

Compiling bus schedules

Approportioning routes and buses to motor pools

Planning the technical and economic parameters of the bus

Service for all levels of the control hierarchy

Controlling traffic on a day-to-day basis.

2. Object Description

The Georgian Republic can boast a highly developed bus service: the annual volume of passenger traffic amounts to 511 million passengers and

* General Director, Production-Technical Amalgamation, G.S.S.R. Dept. of Motor Transportation

4,533 million passenger-kilometers; the republic operates 565 interurban, 1,001 suburban, and 257 urban bus lines, the length of all routes adding up to 108,700 km.

In view of the many random factors affecting the transportation demand, the transport process should be regarded as a probabilistic one. It has been found nevertheless that population movement, for all its probabilistic nature, follows certain laws which shape passenger flows in space and time--and fairly strictly at that. Hence, a possibility arises to work out a network of routes with a regular bus service on the basis of a fixed schedule, thereby converting the probabilistic system into the next best thing to a deterministic one. On the one hand, this approach brings a measure of orderliness into transportation demand as the passengers are enabled to plan in advance the routes and times of travel; on the other hand, however, the passengers' freedom of choice is somewhat curbed. Reconciling these two contradictions is the problem confronting any effort to plan the operation of a passenger transport service as a technical system.

Several problems also arise if the passenger motor transport service is viewed as an economic system. Thus, transportation demand can only be met on the basis of some resources, but these must be administered in the most effective way. Since the available resources are never unlimited and, depending on the particular conditions of use, vary widely in terms of effectiveness of utilization and economics, one finds oneself on the horns of a typical dilemma: either to try to find the most effective use for the available resources, or to plan resources to meet a prescribed level of effectiveness.

Furthermore, serious problems must be dealt with while considering the passenger transport service as a social system. At the present level of civilization, movement turns into one of the most important problems which

may presage the success or failure of any effort to develop productive forces, raise the cultural level and living standards, and meet other social needs of society. Viewed in this light, the passenger transport system is called upon to satisfy transportation requirements on the basis of social needs. Such an approach, however, is fraught with conflict between the demands the transport system must meet as a social system, on the one hand, and as a technico-economic entity, on the other.

Thus, one can conclude that the object under consideration is a large, complex, dynamic and probabilistic system whose planning and control present considerable difficulties.

3. Requirements for an Automatic System for Planning and Controlling the Bus Service

The problems one encounters while working out and adopting decisions to set up, develop, and operate a passenger transport system, are by and large of a poorly structured type. For this reason, no control strategy can rely on a single utility function which would describe the system in all its dimensions: technical, economic, and social. Furthermore, social objectives as often as not defy attempts at quantitative analysis and can only be represented qualitatively. Most technical and economic problems do lend themselves to quantitative description, though some of them cannot be formalized. With this in mind, the requirements to an automatic control system may be stated as follows:

The system must be a man-machine combination, with the man being responsible for the final decision adopted on the basis of a high-order set of objectives which can neither be formalized nor described in quantitative terms

It must conduce to compatible technico-economic indicators for each decision alternative

It must be adaptable to various modifications of invariable and conditionally invariable input information

It must be able to cope with all problems of technico-economic planning and day-to-day control both continuously and discretely, providing final and intermediate results

It must provide all data needed for the purposes of day-to-day control by all hierarchical levels of the ministry's control system, with the information verified as to reliability, completeness, and timeliness, and taking into account the terms of reference of the personnel participating in the planning and controlling of the passenger transport operation;

It must be able to cope with new problems related to the transport system without requiring substantial modification

The above leads to the conclusion that the man-machine system for planning and controlling the operation of the passenger motor transport network must be able to solve problems of two classes: those related to an improved traffic pattern; and those related to the automatic processing of technico-economic and control data.

4. Solution Experience

As an example of the way to solve first-class problems, let us discuss how the interurban bus service of the republic is planned and organized. The traffic plan is based on the population transport demand. Using statistical methods, one can process the data of passenger flow surveys or line documents in order to form a three-dimensional matrix of passenger relationships which would reflect the pattern of population movement in space and time.

A survey of the republican interurban bus service indicates that the passenger flows fluctuate widely both during the day and from one day of the

week to another. Seasonal fluctuations have also been noted. (The survey, conducted in 1970-71, used special questionnaires for each running bus in the interurban bus network filled in every day during one demonstration week every quarter.) After statistical reduction (with a ten-percent deviation from the meanvalues), the data were used to construct a three-dimensional matrix for week days of the same type (working days and rest days) and for different seasons (summer and fall-winter periods). Thus, four combinations of a three-dimensional matrix of passenger flows were obtained, which formed one part of the planning data base.

Another component of this data base is represented by the road network. The peculiar geography of the Georgian Republic shapes its road network which carries interurban buses: almost all towns and population centers are connected by solitary highways and no dense road network is in evidence. So, in actuality, the passenger who wants to travel from one town to another in most cases is left without options. The structure of the highways conforms to a single pattern: four routes of the same configuration sprouting from the republic's capital city. Analysis of the passenger flows suggests that the movement among the population centers lying on different routes is negligibly small and can be ignored to the first approximation. For this reason, a model of the transportation process was constructed on the basis of one route or direction (western), which can be represented as a nondirected graph whose nodes are constituted by the towns and the arcs by the roads interconnecting these towns.

The passenger flows on interurban routes may be carried by rolling stock of various kinds. So, accordingly, the third component of the data base comprises rolling stock characteristics, such as passenger capacity, class

of comfort, and speed of travel over roads of different categories.

The data base is used to plan a regular bus service. The planning procedure consists of three steps, viz. (a) development of a network of routes: (b) assignment of the type of rolling stock; and (c) scheduling of the service. Step (a) implies listing all routes and all buses on each route as well as the time each bus starts to run. Each route is defined by the list of network nodes that the bus must pass in a prescribed sequence in both directions. Assignment of the type of rolling stock implies selecting vehicles with the necessary characteristics corresponding to the principles of the service and permitting to meet them either in the most effective way or at minimum cost. And, finally, scheduling is a procedure of setting the time when the buses must pass all nodes along the route, with the names of the stops and the duration of each stop being spelled out.

The above-described components of the planning process are closely interrelated and thus necessitate simultaneous consideration.

Attempts at presenting all possible alternatives of bus service plans as a system of inequalities and equations generally involve large-dimension problems which often prove too difficult to solve even for computers. So in practice, optimal solutions remain an impossible dream and one has to be content with just rational solutions, at least in some cases.

Therefore, the problem has been solved by heuristic programming techniques. A generalized model of the bus service operation has been constructed on the basis of the following assumptions and constraints which reflect the chief parameters of the passenger movement process:

The policy of the passenger service from the viewpoint of a set schedule does not affect the distribution of the paired relationships in space and time; i.e., the passengers may be classed only as those catered to and those not catered to

The possibility that one and the same vehicle may run on several routes is disregarded

While assigning a bus, the first to be considered are short routes, and then long ones. This sequence ensures a maximum number of through routes and a minimum of transit passengers

Through passengers enjoy priority when buses are assigned to routes, and only then the sections between the various route nodes which are not covered by the shorter routes, come in for consideration

Buses are assigned to routes depending on their capacity, larger-capacity vehicles enjoying priority

A bus run is considered to be acceptable if the dynamic coefficient of bus capacity utilization lies within admissible limits, the lower limit being chosen on the grounds of efficiency while the upper on the grounds of reliability

Permissible limits are set to the time of bus stay at the intermediate stops of the route as well as to the bus driver's working hours

For some route nodes, the time of arrival of the bus at the destination is set at the latest possible hour

The problem can be solved if the output of the system includes the list of routes and the numbers of all buses running on a given route, the time each bus starts its trip, the type of rolling stock, and the values of the dynamic coefficient of capacity utilization. Simultaneously, information is furnished on the bus schedule.

If the above-listed assumptions and constraints are taken into account, the choice of the route network, the assignment of the buses to their respective routes and the selection of the suitable type of rolling stock permit rationalizing the network and optimalizing the rolling stock. For practical

purposes, however, higher-order goals and a number of other constraints introduce certain changes at the expense of system properties.

The above considerations formed the basis of an algorithm which, upon changes in the solution of the problem, forms a new network of routes, selects a suitable type of rolling stock adapted to the new conditions, and calculates the dynamic coefficients of bus capacity utilization. Should the results prove to be unsatisfactory, this algorithm makes it possible to change the initial conditions and perform the next step of iteration until acceptable results are obtained.

Another class of problems aimed at obtaining optimal solutions arises when one considers the possibility that one and the same bus makes several trips. In this case, the number of vehicles for a given network of routes may be minimized. One and the same bus may make two trips provided the following conditions are satisfied:

The end point of one trip serves as the starting point of the other

The time of completion of the first trip coincides with the starting time of the second one (allowing for the duration of the bus stay at the stop)

The overall duration of the two-step operation does not exceed the maximum working hours of the driver (crew)

These conditions lead to mathematical minimization problems which can be solved by known methods.

The bus traffic plan also includes a problem of assigning routes to the various motor pools, the advisability of having each motor pool cater to a specified route depending on the pretrip and posttrip expenses as well as on the technological capability of a given motor pool. The expenses on a zero run and on the maintenance of a bus at the starting point of the trip outside

the motor pool territory (overnight trip) may serve as a criterion for deciding whether or not to assign a given route to a given motor pool. All the other conditions constitute constraints. The solution to this problem yields a list of fixed routes for each motor pool and a matching rolling stock structure.

Thus, the problem of organizing interurban bus services is solved by heuristic modeling techniques involving a certain hierarchy of subproblems, whereby the solution for each subproblem is obtained by descending from optimization to suboptimization or approximate solutions.

After the bus service plan has been compiled, a need arises to determine the economics of the bus operation. Suitable technical-economic indicators characterizing the bus operation may be obtained for each route and for each motor pool on the basis of the bus schedules, the plans of rolling stock assignments to routes, and routes to motor pools.

Aggregation of the route plan indicators by the hierarchical levels of the Ministry of Motor Transport allows for planning passenger traffic for the whole industry in technical and economic terms and objectively enough. It is further envisaged that branch plans will be coordinated with the national economic plans through automatic plan information exchange between the Ministry of Motor Transport and the Republican State Planning committee.

The passenger motor transport system falls in the category of open-loop systems. To make it adaptable to the changing environment, a mechanism is required for disclosing deviations from the plan as well as suggesting and adopting decisions with a view to restoring the balance: i.e., meeting the transportation demand within the set plan targets. Thus, day-to-day traffic control problems must be dealt with. This system must perform the following functions:

Determine the actual state of the system (accounting);

Compare the actual state against the plan and pinpoint all deviations (monitoring)

Analyze the causes behind the deviations (analysis)

Prepare decisions to adapt the system (prediction)

Of the problems listed, the first three readily lend themselves to formalization and can be described by quantitative models, so that automation presents no difficulties here. As for problem four, it belongs to the group of problems which can hardly be formalized. On this basis, when developing the first portion of the day-to-day control system, the primary emphasis was on automating the accounting, monitoring, and analysis functions.

At the present stage, accounting in the bus industry is done by use of three kinds of special documents, viz. driver's trip ticket, bus ticket record, and booking-office record. These documents contain complete information describing the state of the object during one working day, so automatic processing of these data is essential if the accounting problem is to be solved successfully. Currently, the Ministry of Motor Transport is in the process of developing an integrated system for processing the data of transport documents, which, when completed, will register and record the indicators characterizing the operation of all kinds of motor transport. This system will actually mark the emergence of a data bank to cover the entire sphere of motor transport operation. The same system will also permit monitoring the state of any object; i.e., checking the actual data against the planned figures and revealing deviations.

Bearing in mind the multilevel structure of the passenger transport control hierarchy in the Ministry of Motor Transport, effective control requires that each level be supplied with enough information and within a time frame commensurate with the terms of reference of each member of the control system.

To meet this requirement, the system is broken down into hierarchical levels, each having its own list of indicators, sensitivity threshold (the least deviation to which a given level should respond) and schedule of data delivery. Each higher level differs from the one immediately below by an extended list of indicators, a lower sensitivity threshold, and longer intervals of data delivery.

Analysis indicates that quantification of the above components is a formidable task, what with the heuristic nature of the process and the utter impossibility of formalizing the decision formulation and adoption procedures. Hence, the quantitative values of the system components which represent its state were arrived at by means of expert assessment. However, to satisfy the demands of a great number of consumers and make the system quickly adaptable to the individual requirements of managers, the system has a built-in capability of furnishing additional information on demand (a list of questions worked out in advance). Eventually the system will permit a real-time man-machine dialog.

Alongside data on the state of the object, the day-to-day control system must furnish information on the causes of deviation from the planned target figures and also provide initial data for subsequent planning needs. To meet these requirements, special data analysis programs are being developed on the basis of a list of standard causative factors which disrupt the functioning of the system.

With all the elements of the system operating on the basis of the same data, the same software and hardware, we are sure we will be able to provide effective planning and control of the passenger motor transport operation.

AUTOMATION OF BOOKING AND RESERVATION OPERATIONS ON SOVIET RAILROADS

B. E. Marchuk *

It is common practice nowadays to book train tickets in advance through reservation bureaus, transportation and travel agencies, etc.

Acceptance and registration of reservation orders are rather complicated and important procedures. Aside from the sheer number of such orders (hundreds of thousands every day), they come following a highly random pattern, while would-be passengers present extremely diverse demands.

The advance booking system, which leans on a train seat file, does indeed facilitate the booking operation; yet, it falls short of the main objective, viz. high-speed, high-level service.

A typical seat file is found in a special bureau, or in a so-called reservation center which receives telephone requests from personnel manning booking offices and direct calls from passengers.

Under such a reservation system, one dispatcher of the reservation bureau is able to cater to some five permanently operating booking offices. But with a large number of booking offices, this system of dispatcher-controlled operation becomes overcomplicated and all but unwieldy: the number of dispatchers grows proportionately to the number of booking offices; the time of operation gets longer as dispatchers have to queue up if one and the same card is required by several people; finally, longer time is needed for the booking clerk to get through to the dispatcher.

A variety of steps have been taken to improve the dispatcher-control system and thus streamline the train reservation service, but all of them have fallen short of their objective and failed to justify the expenses. The

* Laboratory Chief of All-Union Scientific Research Instutute of Dept. on Railroad Transportation

trouble is that the gain in processing time is disproportionately small if the approach consists in simply beefing up the reservation personnel or increasing the amount of equipment.

There are certain limits to such growth which are rooted in technological, economic, and biological factors, so that at a certain stage of dispatcher system growth the quality of service starts deteriorating: the seat-occupancy rate drops; the newly emergent drawbacks of the system detract from the passenger satisfaction level; the response of the system to ticket cancellation lengthens. This trend is particularly pronounced where the total number of booking workers reaches the 300 to 500 mark.

Hence, a need arose to develop automatic booking systems free from the above-listed disadvantages and capable of effecting a major improvement in the booking procedure.

Past experience suggests that at an annual passenger turnover of up to four million the manual dispatcher booking system is sufficiently effective, but passenger volumes above 10 or 12 million per annum absolutely warrant a switchover to automatic booking.

Automatic booking systems using computers, data transmission facilities, and booking office terminals, provide for automatic collection and processing of all information inputs from travel agencies and booking centers.

Such systems may be of two types: for simple booking, and for the entire scope of booking operations, including automatic sale of all forms of travel documents. Systems of the latter kind, though 5 to 10 percent more expensive, are far more economical, since they operate on the basis of much more extensive passenger traffic information and thus can handle all aspects of the booking operation. The advantages of such a system more than outweigh the extra cost involved.

At present, automatic reservation systems are employed in the railway networks of Japan, Germany, Spain, Italy, France, Denmark and Canada. Of these, the most comprehensive system is Japan's MARS-105 with a capacity of up to a million tickets a day and the capability of catering to some 1,500 booking offices.

Since 1972 the Soviet railways have also had the benefit of an automatic booking system called the "Express" which is superior to the volume of operation to any of its European counterparts. However, automation of booking operations is a far more formidable challenge in the Soviet Union than in Europe considering the large volume of passenger traffic and the much greater railway mileage involved.

The automatic control system "Express" is intended to automate the whole range of booking operations associated with the processes of reserving seats on trains, keeping a check on the seats on long-distance trains, and selling various kinds of travel documents. "Express" is a large computerized control system operating in real time and designed to offer service to a great number of passengers.

Analysis shows that in terms of labor consumption the procedures involved in the booking operation are ranked as follows: ticket registration and documentation of sold tickets with a breakdown by individual booking-clerks, 65%; registration of vacant seats and apportioning them among the booking offices, 20%; and accounting, 15%. It follows that primary emphasis should be on facilitating the booking clerks' job.

Thus, the "Express" automatic control system executes the following functions: (a) checking the vacant seats and furnishing them in response to booking clerks' requests; (b) informing passengers and booking clerks alike as to the availability of vacant seats on trains; (c) determining

the size of the fare; (d) compiling and printing out the various travel and auxiliary documents; (e) computing the amount of money from the sale of tickets and other documents for the whole network of booking offices and for each booking office taken separately; and (f) carrying out all forms of statistical and financial accounting relating to passenger transportation.

Structurally, the system is composed of three elements: (a) a computer center, which is an information storage and processing facility; (b) peripheral (terminal) equipment comprising booking devices and information displays, which are installed at the booking offices; and (c) switchgear and data-transmission equipment linking the computer center with the data terminal equipment.

The computer center is a multiprocessor system made up of three interchangeable computers which control all phases of the booking operation at the various booking points (terminal offices, agencies, railway stations within the city, central reservation bureaus, etc.). Besides, the same computers check and perform statistical reduction of the data on the sale, cancellation, and reservation of train tickets. The computers further compile statistical and running accounts on the utilization of the rolling stock, on the volume of ticket sale for each booking office as a whole and for each clerk in the office, etc. Normally, two of the three computers function in a synchronous duplex mode, carrying out real time processing of the inquiries arriving from the booking-clerks and inquiry devices, while the third computer is in hot reserve, executing statistical accounting functions.

The computer center is so structured as to ensure that the system will continue to function normally even if any two of the three computers fail. The computer center capacity is 200,000 seat orders in 300 trains per day (calculated on the aroung-the-clock basis), with a reservation period of ten days for one-way tickets and forty-five days for return tickets. For people

who wish to reserve tickets long periods (up to two months) in advance, there is an option of acquiring special "Passeord" coupons which guarantee a seat on the train, but the actual tickets (without seat indication) are issued any time before the train's departure.

The booking equipment is designed as an aid to booking clerks and, as such, is installed directly in the booking offices. This set of equipment comprises what is known as booking clerk manipulators (BCM), each consisting of a clerk's console, a ticket printer, and a control device. One BCM can be used by two booking clerks simultaneously.

The clerk's console is built around a full-size keyboard which permits cutting down the time required to set an order and minimizing the error rate. The console is also easy to master and can be readily switched from one function to another depending on where it is installed.

The information on an order being executed may be sequenced as the booking clerk sees fit. This information set includes the following data: point of departure (terminal or station); destination (route); the train number and the type of car which the passenger prefers; the kinds of documents ordered and their quantity; the passenger's privileges, if any; the passenger's requirements to the seat layout; and the type of order (order, cancellation, group order, report, response to an inquiry, invalidation of a wrong document, etc.). If necessary, it is possible to fulfill an order for a specific seat in a specific car.

Booking equipment, depending on the type of job, may be installed both in regular booking offices and in specialized booking offices which do not provide face-to-face service to passengers, such as telephone- and mail-order booking offices; offices invalidating the fares connected with wrong documents issued by booking offices; offices reserving seats on requests from other cities and passing information on vacancies down the train route.

Regular booking offices, which sell and cancel tickets, may be of any kind, from all purpose to narrowly specialized ones. In principle, each booking office could sell and cancel tickets for any train and for any date. The total time it takes the booking clerk to issue two travel documents is not greater than one minute. The maximum daily capacity of a single booking office (with due regard for shift rotation of the personnel) may be as high as 1,500 to 2,000 travel documents.

The ticket printers automatically prepare travel documents on the basis of a unitized letterpress printed form, all travel document forms being letterpress numbered and rigorously accounted for. As a form is fed into the ticket printer, the latter fills in the requisite information obtained from the computer center of the system. The first line of the document produced by the system always contains its name: ticket, pass, children's ticket, excess fare, excess fare to the children's ticket, excess fare to the pass, ticket paid for by written order (check), or servicemen's ticket. If the form is used as an auxiliary document, the first line also carries its name: report, boarding ticket, disembarkation information, etc.

The second and third lines are filled in with the names of the points of departure and destination, respectively. If the travel document is a return ticket, the station names change places.

The fourth line carries digital information (in decimal characters): the number of the booking office that has issued the travel document, the number of the tariff zone of the destination, the time of train departure, the train number, the train category, the car number, the seat number, the ticket price, the document code indicating how the document has been acquired and showing that the advance booking surcharge has been collected.

For passengers planning a change of train en route, the upper right-hand corner of the form bears the names of stations where the passenger is

supposed to pass on the way. The fare is determined by the system automatically on the basis of the points of departure and destination entered into the system by the booking clerk at his console.

Information-inquiry devices exist in several versions: seat availability display boards and information devices for passengers and passenger service personnel. The display board consists of a booking hall board (BHB) and a booking clerk display board (BCDB). The BHBs are to be installed in a prominent position in the large halls of booking offices to inform the passengers as to seat availability. The left-hand portion of each BHB carries permanent information representing the train schedule, while the right-hand side is variable and automatically shows the seat availability picture for each train on the first, second, third, fourth, and fifth day prior to departure. The seat status information for each day is broken down into six categories. The display board is linked with the computer center automatically, so that the information is constantly updated as tickets are bought. Every new day, the information is automatically shifted to the left by the space of one day. A single board is capable of providing information on sixty-three trains.

Booking clerk display boards (BCDB) are installed in the booking clerk booths, enabling the clerk to scan all trains at once for available seats in one of the six categories. The BCDB displays the numbers of all trains which have vacant seats of the category required.

The information-inquiry device is built around a typical ticket printer, differing from the latter only by a simplified manipulator console design and a more versatile printer.

The data-transmission equipment (DTE) may use half-duplex two- or four-wire switching or nonswitching telephone channels, transmission being carried out at a speed of 1,200/600 baud. DTE provides communication links

between the computer center and booking offices as well as between computer centers over distances of up to 5,000 km. The system features thirty-two intercommunication channels and ninety-six channels for communication with the peripheral equipment. The maximum carrying capacity of a single channel is two orders per second.

When the "Express" systems intercommunicate, they share their terminal equipment, for the accessing booking clerk to indicate the number of the particular system he wants. Thus, the booking clerk has access to all the systems and is accordingly capable of procuring seats at all railways. In this case the computer centers which handle the booking clerk's demand operate as information switchboards, storing the data transmitted.

Special switchgear is used to switch demands at the terminal locations.

The experience with the above-described system in the Moscow junction suggests the following conclusions:

1. The booking clerks' productivity can be boosted twofold or even threefold.
2. The booking clerk's duties are largely simplified as the automatic system dispenses with the need to calculate the fare and fill in the travel document forms; nor is it necessary to compile a lengthy report on the travel documents sold, since it is printed out automatically. Under the new conditions, the booking clerk actually becomes an operator who hears a passenger's demand, inputs it into the machine and receives the travel documents printed out by the system.
3. Booking clerks can be trained to work with the system within a short time (one week). The average error rate for booking clerks thus trained is three to four percent.

4. The system marks a dramatic improvement in the quality of service offered to the passengers who are now able to book reservations at any booking office (irrespective of its location), for any direction of travel, and not only personally but also by mail and by telephone. Furthermore, passengers are kept constantly informed about seat availability through information devices.

5. The system raises the cost effectiveness of passenger service through (a) a higher occupancy rate; (b) a more effective utilization of seats as trains move along their routes; and (c) prompt action on the cancellation or use of additional cars or trains judging by the occupancy information (supplied as the tickets are being sold).

6. The system permits improving the quality of planning and evaluating passenger flows on the basis of relevant statistical reports.

According to foreign railways and air carriers, automatic reservation systems conduce to a five to ten percent improvement in the occupancy rate, while the investments into such systems are recouped within three to four years through the cost reduction made possible by higher occupancy rates and boosted profits from advance reservation.

Under current plans, Soviet railways are going to be equipped at the first stage with separate automatic systems to cater to a specified track mileage. As these systems progressively come on line, they will be linked by automatic communication lines, so that eventually all the systems will evolve as a single multicomputer entity with enough scope to cover the entire railway network of the Soviet Union.

THE SIREN SYSTEM: A NATIONWIDE AUTOMATIC CONTROL SYSTEM FOR BOOKING AND RESERVING SEATS ON DOMESTIC AIRLINES

V.A. Zhozhikashvili* et al.

Airline booking and reservation is a typical mass service process of a high degree of complexity steeming from the high frequency of flights and the extremely variable conditions, including the highly fickle weather situation.

That is the reason why air carriers were among the first to employ real-time computerized teleprocessing systems and to have achieved the greatest successes in this field.

Experience of U.S. and European air carriers with systems of this sort suggests that they are instrumental in stemming the tide of personnel expansion, which would otherwise swamp all companies, and in achieving a higher seat occupancy rate on the planes.

Aeroflot differs from most of its foreign counterparts in that it carries out domestic operations on an enormous scale. Thus, the annual passenger turnover is currently approaching the 100 million mark and shows no sign of pace slackening.

The booking procedure for domestic flights radically differs from the international booking operation. Domestic flights have long been a commonplace occurrence, so that it is paramount to relieve the passengers of excessive formalities and make the booking procedure a fast and simple operation. The passenger must be able to book a seat on the plane at any time prior to the takeoff. To achieve this goal, the booking operation must turn into a simple and fast procedure. Besides, domestic flights are far cheaper than international flights. The Aeroflot domestic rates are below

---

* Chief Constructor of the Automated System, Main Computing Center, De[illegible]

those used by most other carriers. Hence, the automatic system should not be overly expensive, calling for a thorough analysis of the various services in economic terms and for a no-frills plan.

These considerations have been taken into account in full measure while developing the Siren system, a unique system in its own right.

The agent-system interaction procedure is extremely simple and fast. In most cases the agent has to access the system only once or twice to cater to each passenger. Tickets are by and large printed out automatically. If need be, the system may automatically print out several tickets, requiring no additional accessing.

The response of the system lies in the range from 1.5 to 3 seconds, while the ticket printing operation takes 5 seconds. Therefore, to take care of a passenger, all the agent needs to do is to receive a demand and the payment, which rarely takes more than one or two minutes.

At this rate, around 20,000 passengers can buy tickets in one hour.

These principles built into the Siren system are going to be extensively used in the nationwide system, too.

The latter system is designed for a capacity of 200 million passengers per annum. No other system in the world can boast such a capacity.

The nationwide system is made up of several zonal subsystems covering the entire territory of the countv. The zonal subsystems include the agencies and airports which territorially fall within each zone. In terms of function, the zonal subsystems are identical and have the same type of configuration. The data processing centers (DPC) are interconnected and linked with their component units by means of a communication network, so that together they form a single system.

The Aeroflot nationwide system was conceived as a single system in the full sense of the word. The centers and zonal subsystems are not going to compete for passengers, but will rather strive for a common goal for the good of a single carrier. While mapping a complicated route involving changes or while searching for an optimal growth strategy, the subsystem centers will view one another as members of the same family pursuing common objectives, rather than as competing or even cooperating companies. This will certainly be reflected in the ways user programs will be executed.

Furthermore, there will be an additional task of optimizing the system as a whole, rendering the Aeroflot system different in essence from the totality of systems run by different carriers.

In technical terms, the system will widely use minicomputers for small centers, alongside huge machines that will be installed in major centers.

The first system, the groundwork of the forthcoming nationwide system, has been in operation since April 1972. It is the above-mentioned Siren which controls over 500 flights from the Moscow airports daily and affords fast and equal access to the seats on these flights for 250 agents in Moscow and another forty-two cities of the country (figure 1).

Designed for a daily capacity of 50,000 seats, the Siren is one of the largest systems in the world in terms of capacity.

In 1972, the Siren gave service to 2 million passengers; in 1973, 4 million; in 1974, 6 million; and the expected figure for 1975 was in the vicinity of 8 to 9 million.

Thanks to this system, all operations are now executed much faster and the productivity of the agents has been raised severalfold, largely adding to the passenger satisfaction index and thereby boosting the traffic volume.

Upon introduction of the Siren, the seat occupancy rate started to grow. In figure 2-a, curve I illustrates the increment of the proportion of

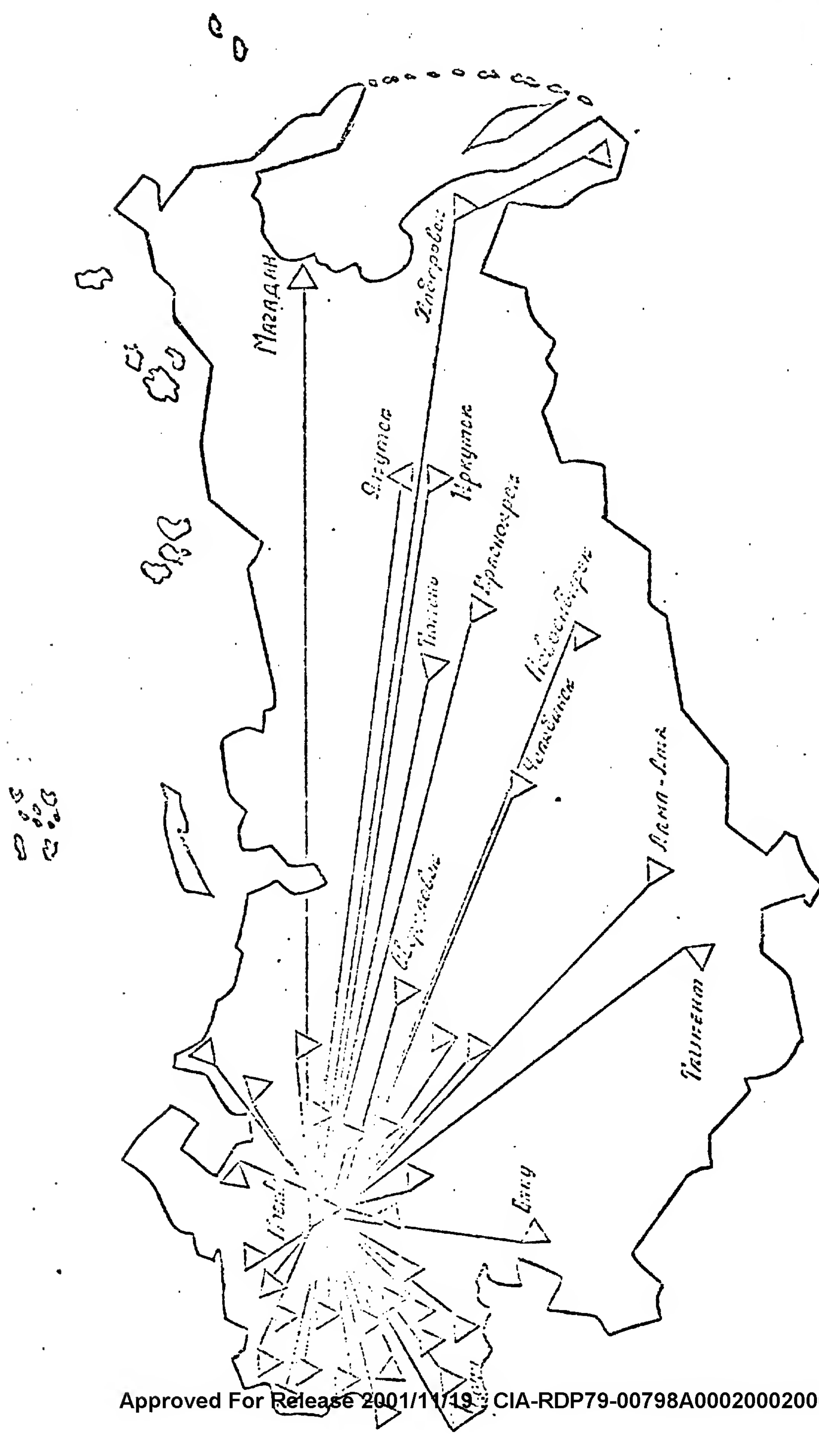


Рис. 1

Figure 1

booked seats on the system-controlled flights, while curve 2 presents a similar picture for the uncontrolled flights. The figure clearly shows that the controlled flights definitely have an edge on the flights operating outside the system. The relative magnitude of this edge is shown in figure 2-b. It is expected that in 1975 it will reach approximately 8%. (The dash lines in the diagrams represent predictions; Ikb means first quarter).

The Siren is built around an integrated system of technical facilities, all made in the USSR., which include a highly automated data processing center, data transmission equipment and a large number of display terminals. The system uses a special language for man-machine dialogue and an operating system controlling real-time multiprogram operations. The Siren is a means of automating a number of services, booking operations, and monitoring the activities of the passenger service personnel.

The need for such an automatic system stems from the fact that the old methods of booking and reservation as well as the techniques for controlling these processes run increasingly counter to the mass traffic requirements in a situation characterized by fast growth of the airline network, larger passenger aircraft, higher speeds, and increasingly hectic pace in the activities of airports and agencies. Thus, prior to the introduction of the Siren, the files of the Main Air Traffic Agency run by the Ministry of Civil Aviation contained up to half a million seat cards which were accessible to over 250 booking clerks, dispatchers of the transit services of the Moscow airports, and personnel of the reservation bureaus and agencies located outside Moscow. Every day up to 8 to 10 thousand seats were booked by cable. In spite of the regularly increasing staff handling the booking demands, the booking clerks had to wait for an average of thirty seconds to one minute, and as often as not up to ten or fifteen minutes, which caused

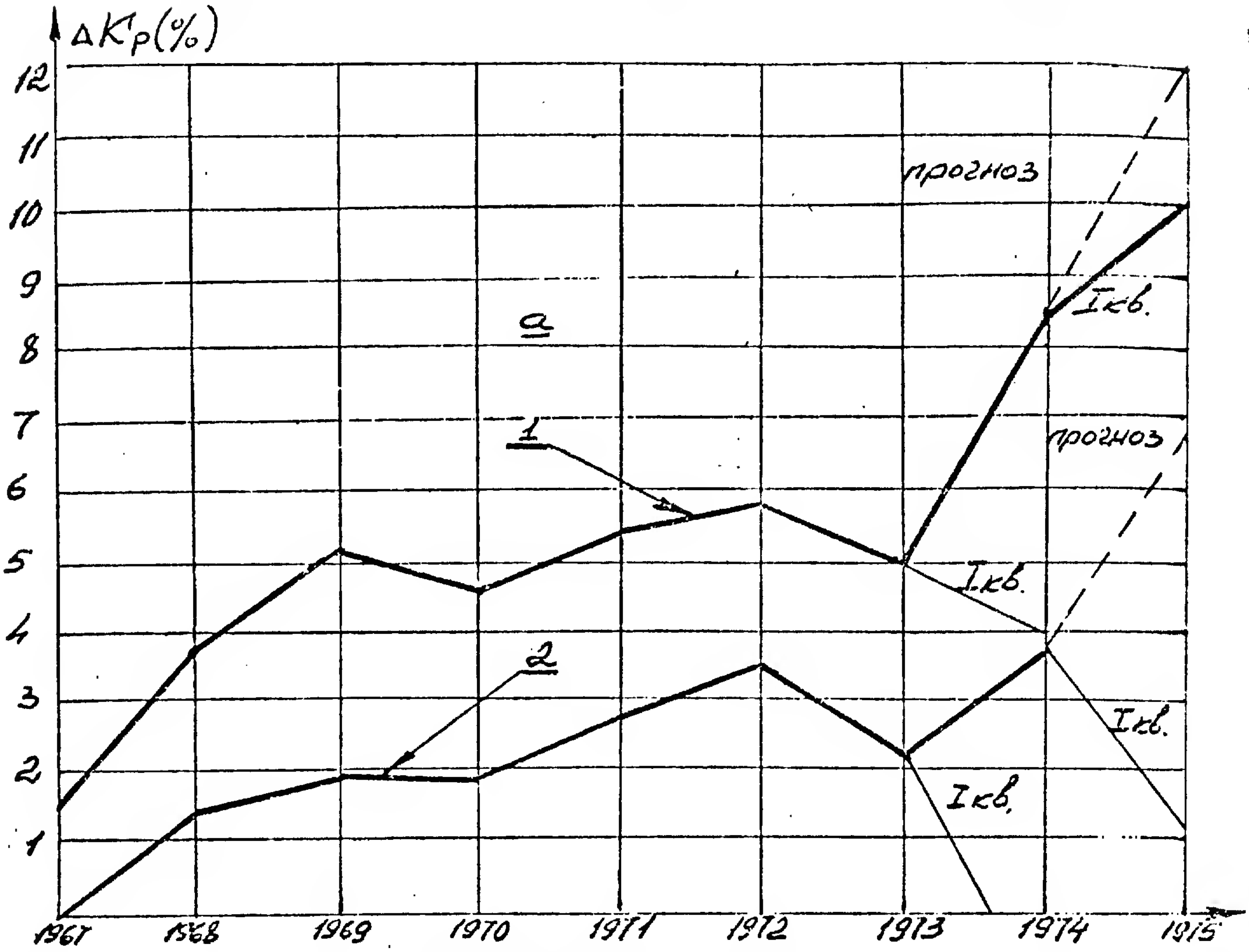
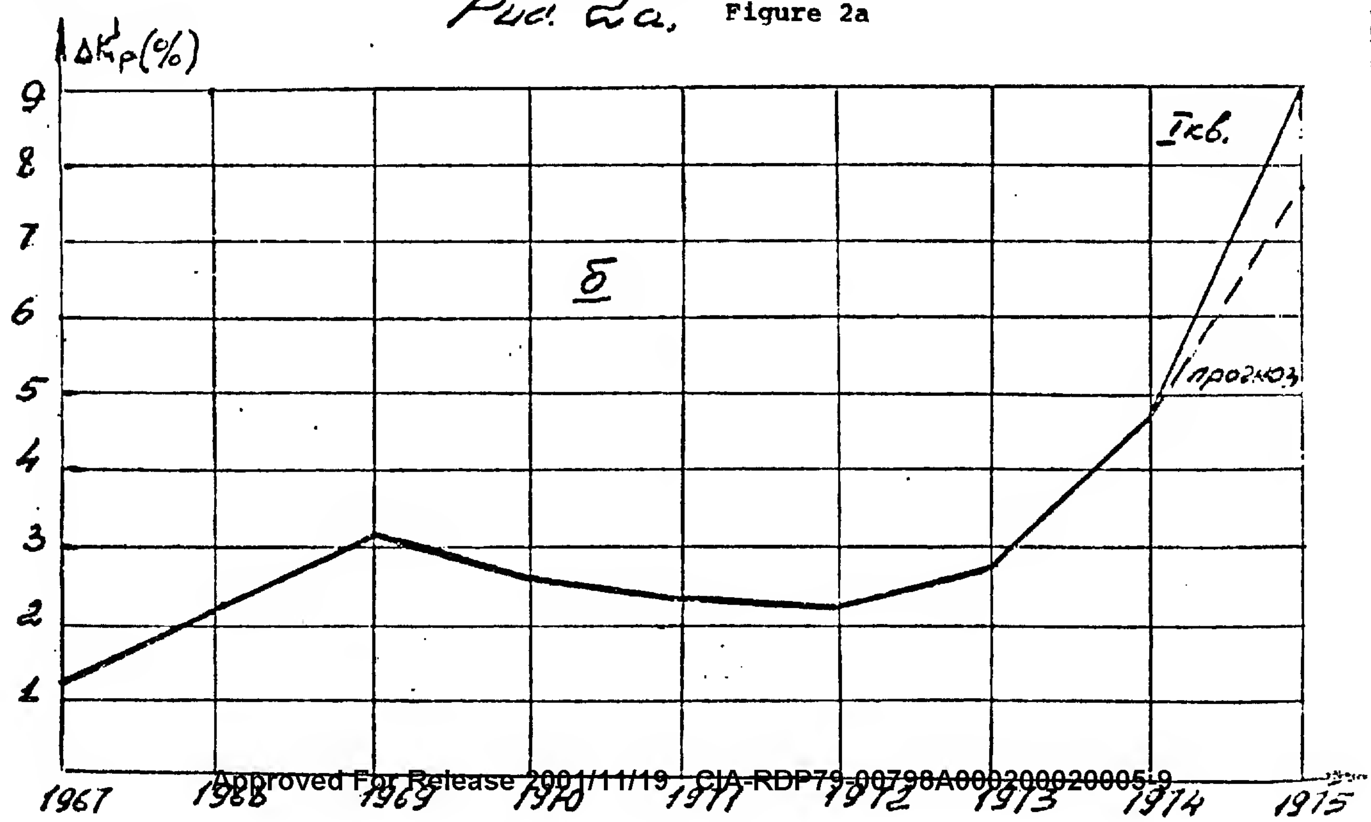


Рис. 2а. Figure 2a



queueing and complaints. The cable demands from other locations for seats for transit passengers were dealt with only the next day.

Curing one shift in Moscow, several tens of thousands of telephone demands would be handled. The hectic pace, the often poor audibility, the inevitable blunting of attention under conditions of incessant monotonous calls, the swarms of people manually writing down short (five to six digits) notes, all added up to errors detracting from the quality of the service and causing much damage. Naturally, execution of orders and monitoring of the personnel's work were all but impossible.

The Siren system automated to a maximum possible degree the most typical operations, furnished the booking clerks and dispatchers with diverse and accurate information and thus made it possible to more than double the booking speed. As a result, the quality of the service was raised considerably and sizeable saving of social time was achieved.

The heart of the Siren is a data processing center (DPC) (figure 3). It is composed of two computer centers with accessory devices whereby the two computers can operate jointly, and special devices interfacing the computer facilities with communication channels. Each computer is equipped with a processor, memory units, input-output devices, and standby facilities.

The DPC memory (both computers) stores all information necessary for the system to perform its technical functions: data on schedules, rates, tariff distances, aircraft seat mock-ups and the sequence in which they are booked, and restrictions connected with traffic rules. The system functions are backed up by a software support system made up of a great number of programs. A specially developed real-time operating system provides for the operation of the DPC in a variety of modes, communication with the many data transmission channels, and response to emergency situations.

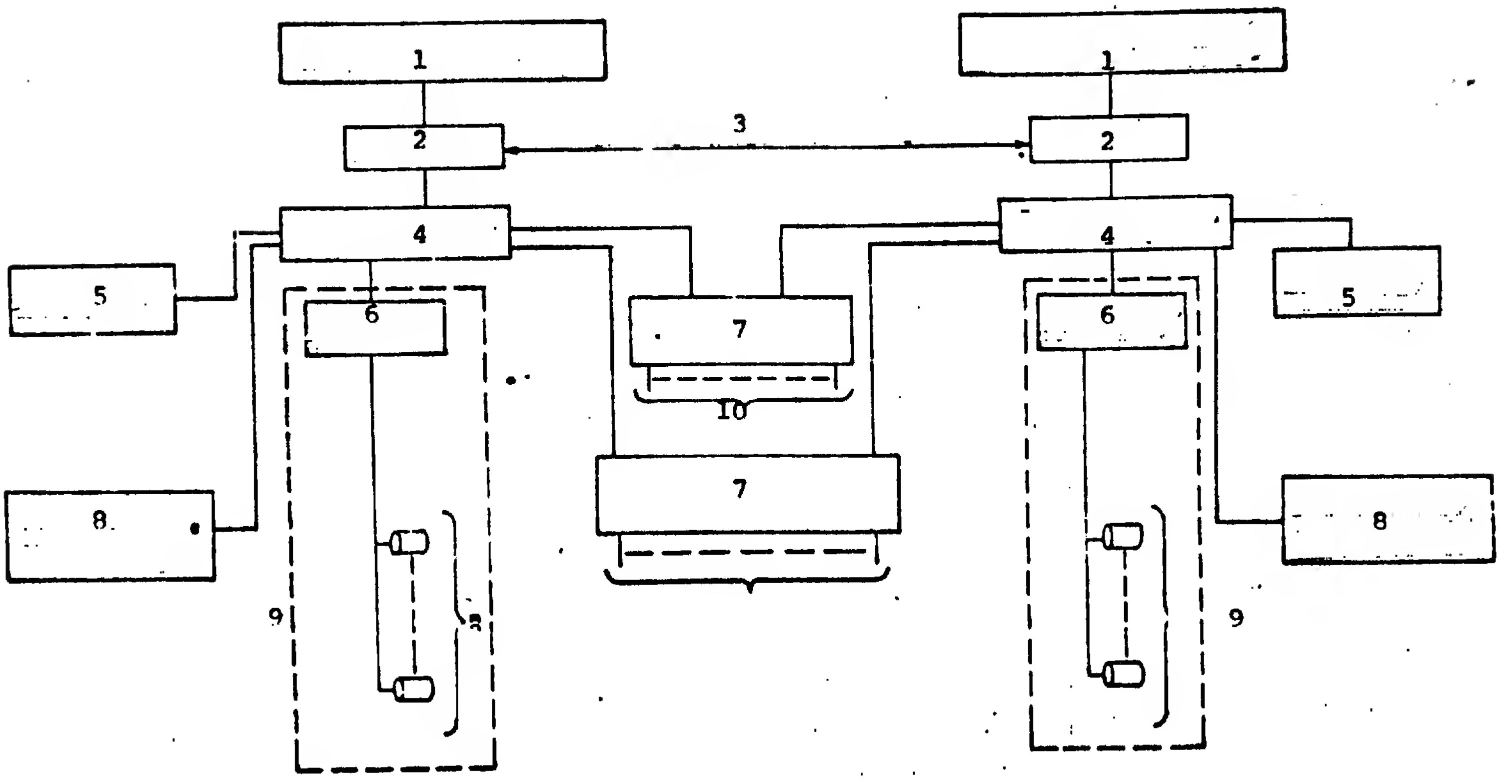


Figure 3: 1 - Mainframe memor (1 reserve)
2 - Processor
3 - Interprocessor link via interface 1B
4 - Input-output channels (1-reserve)
5 - Input/output device
6 - Control devices
7 - Distribution-conversion complex
8 - Magnetic-tape memory unit (2-reserve)
9 - Magnetic-drum memory unit 8 mechanisms (4-reserve)
10 - Communication lines

The chief operating mode of the DPC is duplex operation whereby each request is processed by both computers, the results are compared and issued to the caller in case of complete coincidence of the two sets of results. Should there be a failure of any device, the computer center in which the failure has occurred is automatically switched off, and the system switches into a simplex mode while the DPC operator is notified about the cause of the failure. After the fault has been eliminated, the system reverts to the duplex mode, the switchovers from one mode to the other in no way affecting the processing operation.

The particular configuration of the DPC which involves two independent computer centers capable of functioning both jointly and separately, is an important factor contributing to the realiability of the system as a whole. To the same end, all principal DPC devices have built-in reserve capability and there are special programs for soeedy faultfinding. Thanks to all these facilities, the DPC has a high level of reliability.

Special measures have been taken to protect information. Apart from protection programs, the system uses magnetic drums which continuously record the information status in real time, as well as any change caused by booking, cancellation, or any other operation. This information can be used to restore the system.

The information stored on magnetic drums is updated by the results of each individual access.

Operating experience indicates that at the present level of reliability, the duplex mode is redundant and one of the two computers may be employed to solve concomitant problems in a batch mode.

The main technological unit of the Siren system which provides for man-machine interaction, is a console. The operator's dialog with the system is effected by use of a keyboard and a CRT display which shows the

text of the operator's call and then the information furnished by the DPC in response to the call. The console has a built-in printer for documenting the most important data obtained by the operator. The booking offices covered by the system are equipped with printers for making out tickets.

The Siren consoles are installed in the subdivisions of the Ministry of Civil Aviation Main Agency in Moscow, at the airports of the Moscow area, in some communities in the Moscow region, as well as in the air travel agencies of another forty-two cities of the country. The consoles may perform different functions depending on the particular task with which given subdivisions are charged.

In terms of functions, the Siren system can be subdivided into the following technological units:

Department of short-term planning of commercial flight loads (for flights taking off from Moscow), which includes a day-to-day system control center. The airports of Moscow are equipped with consoles; using these the control center dispatchers execute similar functions, and the system with information about the newly vacant seats immediately prior to flights, and obtain data on the booking operation for transmission to the computer center.

Most of the available consoles are installed in the booking offices which cater to passengers who wish to buy tickets in person. A reservation office has been set up to take care of passengers who make telephone bookings. This office interacts with the system via its consoles and delivers the tickets to the passengers by messenger or else issues them in a special booking outlet at the passengers' convenience.

Certain booking clerks, transit dispatchers at the Moscow airports, reservation agents, and some other workers cannot interact with the Siren system since consoles are not available to everybody. To enable the group

to have access to the system, the personnel of the commercial flight loads division receive telephone calls from them, pass the calls to the system, and transmit the answers to the callers by telephone.

Special mention should be made of a large group of consoles installed in over forty agencies located outside Moscow. Passengers in these communities are able to book seats on flights involving a change in Moscow.

It would be erroneous to assume that the range of the Siren is unlimited by the locations of its consoles. The out-of-Moscow agencies connected to the Siren form their own networks of agencies of sorts in their vicinity, which book the seats on the system-controlled flights through the medium of said agencies with which they keep contact by cable or telephone. Thus, for instance, the consoles of the Pyatigorsk agency cater to all the booking points of the Mineral'nye Vody resort area and the other booking offices in its zone. Similarly, the Alma-Ata agency uses the Siren to book passengers from all cities of Kazakhstan and even some cities outside that republic on flights which involve a change in Moscow. The number of cities thus covered by the system directly or indirectly is well over 400, their network encompassing practically the entire territory of the Soviet Union.

The system contains mock-ups of aircraft with fifty different seat layouts. The seats are sold in two different sequences depending on the type of aircraft, from the front row to the last one or the other way around. However, the system can also accommodate passengers claiming specific seats.

With the information completely centralized and the system operating very fast, the system operators, wherever they may be, enjoy practically simultaneous and unlimited access to all data. Thus, a seat cancelled, say, in Magadan in the far east of the country, may be sold within seconds in Leningrad. It takes from 1.5 to 3 seconds for the answer to come through a telephone channel and 20 to 25 seconds by cable.

The Siren includes an information display showing the situation with seat availability for the next seven days. The display incorporates an automatic calendar and a system-controlled vacant seat indicator. This display is installed on the airport premises in plain view of the passengers.

An important place in the Siren system is occupied by the communication network which links the numerous users with the data processing center. Information is transmitted over standard unswitched telephone and telegraph channels. In order to improve the reliability of the data communication network, use is made of accessory devices, viz, data transmission facilities whose functions include encoding, transmitting, decoding and checking the information being transmitted for correctness.

Console switches, employed in those cases where more than one console is installed, contribute to the effective use of the communication channels. One console switch using one or two communication channels can cater to eight to sixteen consoles.

The system may perform many different functions, such as issuing or cancelling seats and selling tickets; furnishing information about the schedule, seat availability, number of vacant seats on a particular flight, changes in the schedules, etc.

A number of functions of the system are executed requiring no operator's request: calculation of the total rate if a group of passengers requires service; offering the next flight which has the required number of seats; making up records; compiling statistics of, say, unsatisfied demand.

The bulk of requests are concerned with booking.

If the operation is ticket selling, the information fed from the Data Processing Center to the console has the format of a ticket and in the course of printing all the blanks of the form are precisely filled in with the names [illegible] the date and

time of departure, and the rate of payment. The tickets may be sold at a full rate or at a discount (school children, university students, etc.). In the latter case, the system checks whether the particular grade of privilege corresponds to the rate of discount, reminds about the time for booking; if several seats are requested, the system furnishes the overall rate with due regard to the size of discount for each passenger.

The seat issuing operation is executed in those cases where the passenger holds a paid transit or return ticket without a fixed data.

While executing both operations, if the required flight does not offer the needed number of seats, the system automatically retrieves the two closest flights (prior to, and after, the flight in question) which have the number of seats as requested. The system indicates the number of seats available on the flight requested and furnishes all information relating to the two alternatives offered. Thus, the operator is able to help the passenger in choosing the most convenient flight.

The cancelled seats may be returned to the system.

The operator can also use the seat inquiry request to obtain information on seat availability on a particular flight as well as the schedule and rate data. If the seats are unavailable, the system again offers two closest flights. If the passenger has not named the desired flight, but only the desired time of departure, a schedule inquiry is in order. In response to this request the system gives information on several flights in the same direction, the data being sequenced in a chronological order, as well as data on the flight numbers, the airports of departure, the times of departure, and the number of vacant seats on each flight. The package further contains the data on the rate, the day of the week which falls on the date given, and the time of inquiry. In case there are many suitable flights, the inquiry may be narrowed to a certain time interval.

On the operator's demand, the system sums up his activity for the shift or for any time of day, indicating the number of booking operations, seats issued and cancelled, and the receipts.

The dispatchers of the control center run the system through control requests, such as temporary change of schedule for one or several days or permanent schedule change, introduction of a new flight, cancellation of a flight, temporary ban or removal of ban on ticket sales, increase or decrease in the number of seats, substitution of one type of aircraft for another, and the like.

Besides, the dispatcher may request information on the passenger load on any flight, the number of flights in any one direction for which all the seats were sold out, the results of work of any operator or all operators, any subdivision or the whole system, the latter document indicating the number of seats sold as of the time of inquiry, the number of documents issued by the system, the number of operators employed in the system, etc.

All these operations are executed in real time.

The system operators' work is monitored by use of the daily records issued by the system. This document outlines all operations carried out during the day, each operation being entered in the record under the number corresponding to that on the document printed out by the operator and with the time of execution indicated. The records contain information on the type of each operation, the flight number, the date, the seat number, the rate, the kind of discount, and the commission. If the operation was executed outside Moscow, the records will show the date of flight number of the passenger's arrival in Moscow.

Taking into account shift rotation, the Siren system employs over 700 operators, many of whom are located thousands of miles away from Moscow.

With this number, a constant flux of employees is inevitable, which places particular emphasis on the ease of training, on the one hand, and on the foolproofing of the system, on the other. These considerations were allowed for by the system designers.

Information exchange between the operator and the system proceeds in a comprehensible language; the names of cities and many communications are printed out in full; the abbreviations used need no special conversion tables to be understood. The system contains information on a special instruction flight which can be used to train operators on the job.

If the operator does make a mistake, requests information on a non-existent flight, misses some element, or uses a wrong sequence of elements in the set, the input vetting program will block such a request. The screen will display the text of the erroneous request and indicate the nature of the error. Thus, not only is erroneous information barred from the system, but the operator is assisted in locating and easily correcting his mistake. The system, in effect, trains the operators.

Over the period of pilot-scale operation, the system has been enriched with almost forty new programs and changes in the old programs aimed at improving the quality of the service, raising the efficiency and reliability of the system, and facilitating the operators' job. All these changes were effected without shutting down the system.

In 1974, daily control centers were set up at the Moscow airports, marking a new stage in the system development plan. These centers permit tickets to be sold practically until the very takeoff, the seats vacated by passengers switching to earlier flights to be entered into the system, and additional seats to be entered into the system if a flight is executed by a plane with a larger number of seats than envisaged by the schedule. Thanks to these centers, the occupancy rate is bound to rise still higher.

The above-listed advantages of the system can be enjoyed not only by the booking offices of Moscow. The benefits of the system are even more telling in the eyes of the passengers changing flights in Moscow. It must be admitted that the traditional reservation system for transit passengers falls far short of meeting the passengers' requirements. Transmission of cabled requests, processing of cables at the post centers, the booking procedure proper--all take a lot of time, so much so that passengers are able to acquire a transit booking with a confirmed reservation the next day at the earliest, and not infrequently on the third or fourth day. But the Siren revolutionized the entire transit booking business. Now an efficient service, equivalent to that offered to passengers booking tickets in the booking offices of Moscow, is available to thousands of passengers in many cities who fly via Moscow. The operator sitting at a console connected with the data processing center by an ordinary cable link receives information from Moscow within twenty to twenty-five seconds; the information for non-Muscovites is as convenient as that provided to the capital inhabitants. Of course, this modern form of service is in a different class from the traditional cable booking operation. Initially, the plans envisaged that the Siren would cater to twenty-eight cities. However, already today their number has passed the forty mark, and the process of expansion of the system will continue.

It is not in the least accidental that the passenger community has hailed the Siren with enthusiasm: in 1973 the number of bookings from out-of-Moscow agencies increased from 65,000 to 593,000 as against 1972: and in 1974 the volume of bookings skyrocketed by almost three times.

It has been mentioned that the Siren is a man-machine system in which the human operator plays a very important role. The Siren or any other

comparable system calls for an intensive training program (740 operators were trained in Moscow and in other cities), but also for a drastic reconstruction of the organizational pattern, new technology in all segments of the service, new instructions, and new manuals.

In one year, i.e., from May 1972 to April 1973, the Siren gave an economic effect amounting to 3,120,000 rubles, primarily through a higher occupancy rate. The savings of social time constituted about 3 million hours within the same year.

The system of moral and material incentives for the personnel has been revised, and new targets planned for the main agency.

The automatic system also performs a major social function. The operators' condition has been dramatically improved. The booking clerks are now spared nervous breakdowns which earlier accompanied their attempts to place a call to the flight loading people, particularly when audibility was bad; they work much more calmly and smoothly. They can communicate calmly with the passengers and obtain fast and full answers to many questions that would otherwise cause difficulties: they no longer have to consult reference books or seek clarification by telephone. A dramatic cut has been achieved in the number of conflicts because of lack of seats on flights, since the system immediately offers acceptable alternatives, and such an answer obviously meets with a far less adverse reaction on the part of the passenger. The passengers also feel much better because now they have to wait in line half the time they would spend under a manual booking system, even during periods of peak demand.

The situation has changed even more dramatically for the better in out-of-Moscow agencies equipped with consoles. Thus, while earlier passengers had to come at least twice to procure a seat on a flight passing through Moscow, now they receive their tickets immediately. This affected in a major

way the structure of transit passenger flows: the number of passengers coming to Moscow on an "open" ticket, i.e., one containing no date of the next flight, has been cut down considerably. The vastly facilitated reservation procedure has boosted the number of people who prefer to fly via Moscow, having previously reserved a seat on a specified outbound flight.

As transit passengers reserve seats with a change in Moscow, they are immediately told the seat number, which relieves them from the duty of having the ticket registered by the transit dispatcher and thus spares them the formerly unavoidable exhausting waits to obtain a seat on the next leg of their travel. This has also benefited the transit dispatchers of the Moscow airports.

For all these reasons, the volume of passenger traffic has risen by a substantial margin. Thus, during the first ten months of 1974, the seat occupancy rate for the Moscow air junction rose by 3.1% as compared to the comparable period of the previous year, and amounted to three times the average figure for the rest of the Aeroflot units.

The experience with the development and introduction of an automatic system for booking and reservation operations has proved beyond all doubt that systems of this type can find wide application as a means of automating many processes in the civil aviation field which are characterized by high speeds of execution and call for fast decisions within a limited time. Examples of such operations are as follows: daily control of airports, short-term crew planning, redistribution and mapping of optimal routes for supplying spare parts to grounded planes, inquiry and information service, information flow control on a nationwide basis, and many others.

Centralized storage of vast quantities of diverse and rapidly changing data, availability of this information at any point in time to any user who may be thousands of miles away from the source of data, centralized storage

and allocation in real time of limited resources (plane tickets, hotel rooms, hospital beds, tourist passes and the like) to hundreds of users, high reliability of information, high degree of automation of operations--all these factors testify to the enormous potential of teleautomatic mass service systems, such as the Siren, and not only in civil aviation, but in many other branches of the national economy as well.

The large body of designers working with the Ministry of Instruments, Means of Automation and Control Systems, together with the experts of the Ministry of Civil Aviation, who designed and built the Siren, can be justly proud of a major achievement in developing an automatic mass-service real-time system with a unique information network, which links the remotest cities of the country with the central computer in Moscow.

INFORMATION MODELING FOR MANAGING SOVIET MARITIME TRANSPORTATION

V. S. Bondarenko *

Goals and Methods of Modeling

It is common knowledge that the control process is composed of the following elements: acquisition of initial data characterizing the status and behaviour of the object of control and of the environment; transmission of these data to a control facility; accumulation, storage, and analysis of the information thus received by the control facility; formulation of possible alternatives of control decisions; selection of the best decision for a given set of circumstances; and transmission of the latter decision to the object of control.

Formulation and selection of control decisions are procedures which lend themselves to varying degrees of formalization, may or may not include optimization algorithms, and may or may not be computerized. However, whatever the case may be, no control decision can be developed without initial information which is liable to vary widely in terms of composition, contents, acquisition intervals, and other parameters, the variation being due to both objective causes (different objects of control, different functional goals of the objects of control, different quality criteria and the like) and many subjective factors, including individual traits of the control facility managers and rank-and-file personnel.

Since the real objects, their functional processes and control algorithms, are highly complex, a huge number of control alternatives and for practical purposes, an infinite number of information models can be suggested. Each information model features its own depth and degree of description of the object of control, its own mode of inputing data into the control system, and

* Chief of Main Computing Center, Department of Sea Transportation

its own quantitative and qualitative parameters characterizing the control system's response to changes in the object of control and in the environment.

In order to provide the most rational type of control systems combining high efficiency with favorable economics, it is paramount to analyze thoroughly and prove the all-round advantages of a suitable information input.

The comparative approach to the evaluation of various input systems and the fact that some of the possible alternatives simply cannot be realized, suggest that it may be helpful to try and simulate information processes.

Information modeling in this context is viewed as an organic component of the comprehensive control system model, on a par with economico-mathematical modeling of objects of control where problems of optimization, prediction, and analysis are involved.

However, as distinct from economico-mathematical modeling employed in highly developed control systems which are called upon to formalize the procedures of optimal decision making and quantitative analysis of trends, cycles and future programs, information modeling is feasible in all kinds of control systems, including the most trivial computerized and noncomputerized data processing systems.

Information models can be broadly classified into two types: particular models characterizing some local or narrowly functional elements of the information input of real control systems, and generalized models which may be regarded as analogs of full inputs of real control systems.

The matrix techniques of analyzing and synthesizing the information inputs of control systems, which are widely used in modern design institutions, can be successfully employed for constructing particular and generalized information models. Comparative technico-economic analysis of these models--particular and generalized alike--is a method for subsequently selecting decisions most conforming to the goals and objectives of control.

The importance of information models rises as control systems develop, i.e., with the degree of their sophistication, resolving power and, naturally, cost.

Objects of Simulation

The information models of control systems should reflect, within certain limits of accuracy, the state and behaviour of the objects of control and of the environment, on the one hand; and the response of the control systems to the changes occurring in the objects of control and in the environment, on the other. The groups of indicators or information modules which define the state of the object of control and of the environment, present an important characteristic of the information input of any control system and, hence, must be among the primary parameters coming up for simulation study. The most efficient set of indicators is apparently one that is minimal but sufficient for describing the object of control and its environment in terms of formal and informal decision-making algorithms used in a given control system. But the algorithm structure, in its turn, depends on the available data base as well as on the scope of numerical methods and the technology. So, modeling of the set of indicators is closely associated with the economico-mathematical modeling of the object of control and decision-making processes. If all these models are viewed together, it is possible to select effectively the important characteristics of the object of control for regular measurement, registration, and inputting into the control system.

The need and advisability of inputting into the system of a group of indicators characterizing the state of the object of control and of the environment depends on a number of other factors:

1. Forecast of changed requirements to the data base of the control system brought on by the evolution of the object of control and by the dynamics of its functions and control objectives
2. The technological capability of the available facilities for data acquisition, transmission, and processing as well as the prospects of their development
3. Economic considerations and, among other things, evaluation of the costs of increasing the amount of information inputted into the system against the effect due to the possible improvement of the quality of control

In multilevel control systems, the number of hierarchical levels and the way information-computation tasks are allotted to them may affect the technical characteristics and the economics of the system. Consequently, information modeling may be set a goal of determining the most efficient indicators of information filtration and condensation as the data move from one hierarchical level to another. Since such indicators depend on the ability of each level to process information and make decisions, analysis of the various patterns of allotment of informational and computational tasks in the course of information modeling may lead to an efficient vertical structure of control facilities and the optimal information volumes to be processed on each level.

The modes of inputting and processing of the starting data are important functions of real control systems, which by and large determine the information capacity and sensitivity of the systems. The frequency at which the on-line numerical values of the parameters of the object of control and of the environment are registered and introduced into the system, is either determined by the control resolution (critical value of displacement) or else regulated by preset time intervals. In the course of information

modeling, the mode of information registration and inputting is varied and each alternative evaluated in technical and economic terms in order to make a justified choice of an optimal or near-optimal "pulse" of the control system.

To combat irregularities in the arrival and processing of information in the course of information modeling, the data are smoothed through varying the information density of the various groups of indicators during the week or the day, taking into account the effect of information delays on the performance of the control system.

Modeling of data arrays and banks gives an idea about the structure and configuration of the data base which is most advantageous for a given control system. Bearing in mind that the data base is among the most crucial elements of modern control systems, the various configurations of the arrays making up the data bank as well as the various relationships between them are analyzed by comparing the costs of setting up and operating data banks atainst the current and predicted requirements of the control systems with a view to avoiding any reconstruction costs in the future and minimizing the direct and indirect expenditures.

Any general decision offers some advantages as well as some disadvantages as against a totality of particular decisions. Hence the need to compare generalized data banks with specialized arrays before deciding which alternative to adopt. The technical parameters and the economics of a control system are known to depend on how successfully one evaluated the comparative advantages of slow and fast external memory units, determines the efficient balance between the volumes of initial data and the intermediate information package resulting from statistical generalization and smoothing (mean values, indices, characteristics of dynamic series and the like), and evaluates the various data configurations and relationships between arrays.

Summing up, one can assert that the goal of information modeling is to provide answers to the following questions:

1. What kind of information is needed for effective control?
2. What is the economic time frame of data acquisition, transmission and processing?
3. What are the requirements to the quality of information: completeness, reliability and timeliness?
4. How should the information input develop to match the evolution of the object of control and of the control system?

An optimal data base for a control system may be put together by aggregating the partially optimized decisions on the composition of the set of information indicators, the allotment of informational and computational tasks to the various levels of the control hierarchy, the frequency with which information modules are entered into the system, the intervals at which control decisions are made, the success in smoothing irregularities in the system operation, as well as the composition and configuration of the data arrays and banks.

Modeling Experience

Information modeling is employed for developing an automatic system for controlling sea transport operations. In this system, the chief information elements are shipment orders and data characterizing the operation of ships, the situation on routes and in ports, as well as the international cargo and freight setup. An important role is also played by a constant flow of information on the cargo shipment capability, technical and operating characteristics of ships and ports, distances between world ports, freight and tariff rates, as well as other information of a technical and economic nature stored in the system.

While designing the automatic control system, the indicators included in shipment orders, ship and port records, freightage documents, as well as all the other indicators contained by permanent and dynamic arrays, were selected by the results of an all-round analysis of appropriate information models. The models were obtained on the basis of currently available algorithms for optimizing control decisions, with due regard for the requirements and wishes of the managerial and rank-and-file personnel of the control bodies as well as for the recommendations of the designers of the functional subsystems incorporated in the automatic control system under development.

The structure of the functional subsystems was defined by the results of analysis of the activities carried out by the sea transport control bodies. It included:

- a freight market analysis subsystem, which uses standard freightage reports as the input, and the results of analysis and forecasting of the freight market in its various parts by the kinds of cargo and by season as well as other freightage conditions as the output;

1. A subsystem for day-to-day accounting and analysis of the merchant marine and port operations. In this subsystem, the input is composed of daily reports of ships and ports, plans, schedules and timetables as well as various operating standards; whereas the output consists of reference and analytical material characterizing the state and dynamics of shipment operations and the data describing how the plans, schedules, and timetables are being fulfilled in terms of fleet operation and port work, and how the operating standards are being compiled with.

2. A statistics and bookkeeping subsystem, in which the input is built around statistical and bookkeeping reports of the shipping lines

and ports, while the output comprises analytical material characterizing the financial status of the merchant marine and its subdivisions with a breakdown into periods and operations.

3. A subsystem of short-term planning of fleet and port operation. Its input includes the cargo owners' orders, operating standards, information on the freight market setup and on the times and places of the ships' fulfilling their obligations. The output of the subsystem consists of optimal draft plans of fleet utilization as well as the schedules and timetables of ship operation.

4. A subsystem of long-term planning of sea transport development. The input of this subsystem comprises the starting material for forecasting shipping operations and for producing economico-mathematical models of the optimal strategy of development and allocation of the sea transport resources. Its output consists of the recommendations on plan and design decisions.

Alongside these principal subsystems, the sea transport control system being developed includes three back-up subsystems: technology monitoring subsystem, logistics subsystem, and personnel subsystem.

Analysis of variants was found to be the simplest and most effective means of studying information models and selecting the best design decisions. The variants differed in the comprehensiveness of the information picture and in the degree to which they met the demands and wishes of the control personnel. A list of compulsory indicators was compiled, which included all factors indispensable for making forecasts and plans of shipment, ship schedules, and other regular control decisions, as well as for systems analysis of the most important aspects of sea transport operations.

Those indicators which are only needed from time to time and only in response to special requirements or to a disruption of the normal course of shipment, are considered as claimants to be included among the compulsory indicators in case statistical evaluation shows them to be needed frequently enough and economic assessment indicates that the benefits derived from their routine acquisition outweigh the added costs.

The final selection of the claimant indicators was made on the basis of several criteria, the most important of which were control problems for which they were considered necessary, the frequency of their active use, and the costs involved in their acquisition, transmission, and storage.

To determine the relative contributions of the indicators into the control process, use was made of questionnaires distributed among the designers of the functional subsystems and the control personnel. Also used were matrix and statistical methods of analysis. The results of this work went into a set of indicators made up of several different groups which were to be routinely inputted into the system.

The frequency of inputting the various groups of indicators was worked out on the basis of economic considerations, taking into account the impact of the inputting intervals on the control system performance. Since direct relationships were impossible to obtain, the method used was expert assessment subsequently subjected to statistical evaluation. Where input data serve to obtain mean values, such as characteristics of a dynamic series, standards and the like, an additional factor was considered, viz. the stability of the mean values and the stability of the results of control problems derived on the basis of said mean values.

The three-level structure of the control hierarchy (shipping line-agency-ministry) mandated a study into the quantity of information needed by all levels to cope with their respective functions. Here situation

simulation and analysis of alternatives were used. As a result, decisions were made whereby the second and third control levels jointly served by the Main Computer Center were to be furnished with substantially smaller quantities of information than previously envisaged, and the information reaching these levels was to undergo special screening.

All main information arrays were designed as generalized entities to cater simultaneously to many subsystems and many control problems. The standardized structure of the information arrays allows of a gradual switchover to a systematically organized data bank.

Apart from a considerable economy of inputting and updating costs, such a data bank makes it possible to automate the programmed printing-out and displaying operations. Thus, the new demands are to be presented in a specially developed high-level language whose vocabulary is unambiguously interpreted by the standard program modules. The language vocabulary was composed of the names of standard arrays and their elements, and the translation and editing procedures were tied in with information elements and formally described together with the latter, dramatically cutting down the programming time and improving the performance of the entire control system.

Modeling Effect

Information models are an effective means of evaluating different configurations of the control system data base and choosing a good decision before the stage of actual design, thus saving a lot of time and money when designing new, and reconstructing installed, control systems and--which is particularly important--affording a practicable way of optimizing control systems.

Information alternatives are evaluated on the basis of the following factors: the amount of money spent on designing and introducing a given

system or its elements featuring given information characteristics; average annual costs of maintaining the system or its elements; and average annual costs of progressive modernization of the system or its elements.

Unfortunately, the information models can be compared only on the basis of scant and unreliable initial data. Here the basic principles were as follows:

1. As the accuracy of the initial data decreases, the methods of comparison and evaluation become rougher.
2. Particular decisions require more detailed calculation and more precise techniques of comparison.
3. If reliable initial data are lacking, it is mportant to obtain as many expert evaluations as possible.

The results of evaluation of the effect of information modeling indicate that modernization costs are substantially lower than the overall costs of control system design operations, and modeling effects such savings that the modeling costs are recouped many times over within a few years of the system's operation. It is also clear that if the objects of control are to function efficiently, the inner structure of the control system must be optimized.

Information modeling is not to be considered as an independent stage of design work. It must be carried out in the course of work on the draft in all cases where the control system being developed is complex enough to warrant a search for a priori rational design decisions.

UTILIZATION OF COMPUTERS FOR THE MANAGEMENT AND DEVELOPMENT OF TRANSPORTATION SYSTEMS

Approved For Release 2001/11/19 : CIA-RDP79-00798A000200020005-9